



601061-ME-Memoria II

# **Scoring Proactivo para Particulares**

## **Memoria II**

### **Segmento Activo**

**BBVA-Continental**

## Control de versiones

Versión	Fecha	Revisión	Comentarios
001.000	04/11/2015	Júlia Densalat, Rossana Giaconi	Versión Inicial
002.000	27/11/2015	Júlia Densalat, Rossana Giaconi	Se añaden los apartados 2.4.1.6 y 3.3. Se actualizan las tablas de eficiencia del apartado 4.2  Se agregan los resultados de eficiencia del modelo en los Colectivos de Control en el apartado de Resultados
003.000	04/01/2016	Júlia Densalat, Rossana Giaconi	Se cambian las variables de comportamiento del modelo Proactivo a solicitud de BBVA Continental y se asignan pesos expertos a la nuevas variables. A raíz de este cambio, se actualizan todos los resultados referentes al modelo.
004.000	25/02/2016	Júlia Densalat, Rossana Giaconi	Se completa la definición del público objetivo del segmento Clientes de Activo.
004.001	19/05/2016	Júlia Densalat, Rossana Giaconi	Modificación del formato de algunas tablas para evitar que se corte el contenido de las mismas

# Índice

<b>1</b>	<b>Introducción .....</b>	<b>5</b>
<b>2</b>	<b>Metodología de desarrollo y validación .....</b>	<b>6</b>
2.1	Características de la muestra .....	6
2.2	Definición del evento a predecir.....	7
2.3	Tratamiento de los datos .....	7
2.3.1	<i>Construcción de las variables explicativas (WOE).....</i>	<i>7</i>
2.4	Metodología de Modelización .....	8
2.4.1	<i>Selección del modelo estadístico .....</i>	<i>9</i>
2.4.2	<i>Validación del algoritmo .....</i>	<i>14</i>
<b>3</b>	<b>Descripción del modelo .....</b>	<b>16</b>
3.1	Modelo Proactivo .....	16
3.1.1	<i>Información de Cabecera .....</i>	<i>16</i>
3.1.2	<i>Información de Banca por Internet .....</i>	<i>17</i>
3.1.3	<i>Información de Pasivos .....</i>	<i>17</i>
3.1.4	<i>Información de Tarjeta de Crédito.....</i>	<i>19</i>
3.1.5	<i>Información de Préstamos.....</i>	<i>20</i>
3.1.6	<i>Información de RCC .....</i>	<i>22</i>
3.1.7	<i>Información de Comportamiento .....</i>	<i>23</i>
3.1.8	<i>Obtención de la puntuación del Modelo Proactivo .....</i>	<i>24</i>
3.2	Integración Proactivo + Grupo Riesgo Buró .....	24
3.2.1	<i>w4_cod_crriesgo- Grupo Buró.....</i>	<i>25</i>
3.2.2	<i>Obtención de la puntuación final del cliente .....</i>	<i>25</i>
3.3	Aspectos cualitativos del Modelo Proactivo.....	26
<b>4</b>	<b>Recuento de Resultados.....</b>	<b>27</b>
4.1	Ranking de las variables del modelo .....	27
4.1.1	<i>Ranking de las variables del modelo Proactivo.....</i>	<i>27</i>
4.1.2	<i>Ranking de las variables del modelo integración Proactivo + Grupo Riesgo Buró ..</i>	<i>28</i>
4.2	Eficiencia.....	29
4.2.1	<i>Modelo Proactivo .....</i>	<i>29</i>
4.2.2	<i>Modelo integración Proactivo + Grupo Riesgo Buró .....</i>	<i>31</i>
4.2.3	<i>Colectivos de Control.....</i>	<i>32</i>
<b>5</b>	<b>Validación.....</b>	<b>33</b>
<b>6</b>	<b>Anexo: gráficos woe y reporte técnico .....</b>	<b>34</b>
6.1	Woe de las variables del modelo .....	34
6.1.1	<i>Cabecera .....</i>	<i>34</i>
6.1.2	<i>Banca por Internet .....</i>	<i>36</i>
6.1.3	<i>Pasivo .....</i>	<i>37</i>

6.1.4	<i>Tarjeta de Crédito</i> .....	40
6.1.5	<i>Préstamos</i> .....	42
6.1.6	<i>RCC</i> .....	43
6.1.7	<i>Comportamiento</i> .....	45
6.1.8	<i>BURÓ</i> .....	47
6.2	Reporte técnico del modelo integración Proactivo + Grupo Riesgo Buró .....	48

---

# 1 Introducción

---

El presente documento contiene la especificación del modelo Scoring Proactivo de Particulares desarrollado por AIS para la cartera de clientes de Activo de BBVA Continental. Dicho modelo corresponde a un modelo estadístico desarrollado sobre la población de Clientes de Activo, es decir, clientes sin domiciliación de Pago de Haberes, que tienen algún producto de Activo vigente en la entidad a fecha de referencia. En el caso de que un cliente tenga exclusivamente una tarjeta de crédito en la entidad, éste será considerado cliente de Activo si ha usado la tarjeta alguna vez en los últimos 6 meses, en caso contrario, el cliente no será considerado cliente de Activo. Dicho segmento corresponde a uno de los cuatro segmentos en los que se ha decidido segmentar la cartera de clientes Personas Naturales en su vertiente de Particular de Banco BBVA-Continental.

Además de la especificación del modelo Scoring Proactivo, este documento contiene la especificación de la integración de dicho modelo con el modelo de Buró disponible actualmente en la entidad. Para ello, se ha utilizado una variable proporcionada por BBVA-Continental que indica el Grupo de Riesgo en el que se encuentra el cliente según la puntuación del modelo de Buró.

El documento se ha estructurado del siguiente modo:

- Metodología de desarrollo y validación del algoritmo, descripción del tratamiento realizado para obtener el algoritmo y de las diversas validaciones realizadas.
- Descripción del algoritmo del modelo Proactivo y de la integración de éste con el modelo de Buró actual.
- Recuento de resultados, donde se realiza una valoración de la bondad de ajuste del modelo.
- Validación del algoritmo.
- Anexo: Gráficos WOE y reporte técnico de la integración del modelo.

## 2 Metodología de desarrollo y validación

Para la obtención del modelo se realizan los siguientes pasos:

1. Segmentación



Cuyos resultados se muestran en el documento de Segmentación

2. Definición de muestras

3. Análisis estadístico - descriptivo de la misma



Cuyos resultados se muestran en la *Memoria I*

4. Obtención del algoritmo propiamente dicho

5. Validación del Modelo



Cuyos resultados se muestran en esta *Memoria II*

Próximos pasos:

6. Cálculo de Límites



Cuyos resultados se mostrarán en la *Memoria III*

### 2.1 Características de la muestra

El tratamiento de la información es la clave en el desarrollo del algoritmo. Así, para ser utilizable, la muestra debe ser representativa estadísticamente y contener información de calidad suficiente.

En el caso que nos ocupa, la muestra seleccionada consta de 203,577 clientes no repetidos, 26,554 de los cuales resultan ser morosos, por lo que la tasa de mora es de un 13.04%. Dichos clientes fueron extraídos de tres cosechas diferentes con el fin de mitigar los efectos de la estacionalidad de algunas variables. Para ver la selección de la muestra, consultar el documento AIS-601061-ME-Memoria I-ProactivoParticulares\_SG4.

La muestra ha sido dividida aleatoriamente en:

- 80% (163,082 clientes) **muestra de desarrollo** con la que se ha construido el modelo
- 20% (40,495 clientes) **muestra de validación** para poder aplicar así la técnica de validación del algoritmo “out of sample” explicada en el apartado 2.4.2

## 2.2 Definición del evento a predecir

---

Se define como variable a explicar, la mora (o default) en los 12 meses siguientes a la fecha de evaluación del cliente. Se considera que un cliente es moroso si existe algún impago, en cualquiera de sus productos de activo contratados en BBVA Continental, superior o igual a 60 días en el período de 12 meses de observación de la mora. Así, la variable a predecir es una variable dicotómica que identifica con 1 a los clientes morosos y con 0 a los no morosos.

## 2.3 Tratamiento de los datos

---

El tratamiento de los datos consiste en transformar las variables originales con el objetivo de obtener variables artificiales que tengan un alto poder discriminante. Esta transformación da lugar a los siguientes tipos de variables:

- **Variables agrupadas**, resultantes de agrupar variables categóricas originales con un número muy elevado de categorías o con algunas de ellas con pocos casos.
- **Variables tramadas**, resultantes de realizar tramos sobre las variables numéricas originales.
- **Variables artificiales o calculadas**, definidas a partir de las variables originales. En este grupo se encuentran, entre otras, combinaciones entre diferentes variables.

En la Memoria I del proyecto (AIS-601061-ME-Memoria I-ProactivoParticulares\_SG4) se expone el análisis descriptivo del universo inicial de variables candidatas a formar parte del modelo. Para cada variable presentada en dicho documento, se realizó un análisis bivalente con el objetivo de estudiar así, el comportamiento de cada una de las variables frente a la morosidad y frente al Weight Of Evidence (WOE). A partir de la Memoria I, se seleccionaron las variables con mayor potencial para ser consideradas a priori como variables explicativas del modelo.

Durante el proceso de modelización, se han creado nuevas variables artificiales no incluidas en la Memoria I, por este motivo, en el apartado 6 de este documento se muestran los gráficos bivariantes de todas las variables que participan en el modelo.

Una vez seleccionadas las variables que a priori son posibles candidatas a entrar en el modelo, se inicia el proceso de construcción de las variables que se considerarán como variables explicativas del algoritmo de Scoring.

### 2.3.1 Construcción de las variables explicativas (WOE)

---

Las variables candidatas a ser variables explicativas del modelo se construyen calculando el *Weight of Evidence (WOE)* de cada atributo o grupo de atributos de las diferentes variables.

Este valor es una medida entre la diferencia de las proporciones de buenos totales de la muestra y malos totales de la muestra que se encuentran en cada atributo (agrupación de categorías) de la variable.

La definición del WOE es la siguiente:

$$WOE_{ij} = 100 \cdot \ln \frac{\text{Distribución Buenos del atributo } j \text{ de la variable } i}{\text{Distribución Malos del atributo } j \text{ de la variable } i}$$

Donde  $WOE_{ij}$ : se define como el valor WOE del atributo  $j$  de la variable  $i$

*Distribución Buenos del atributo  $j$  de la variable  $i$  se define como:*

$$\frac{\# \text{ Buenos en el atributo } j \text{ de la variable } i}{\# \text{ Buenos de la variable } i}$$

*Distribución Malos del atributo  $j$  de la variable  $i$  se define como:*

$$\frac{\# \text{ Malos en el atributo } j \text{ de la variable } i}{\# \text{ Malos de la variable } i}$$

Una vez se ha tratado la información tal como se ha comentado y se dispone de las variables descritas, se inicia el proceso de construcción del algoritmo de Scoring a través de Regresión Logística.

## 2.4 Metodología de Modelización

Para variables dependientes o de respuesta binomialmente distribuidas, la regresión que se aplica es la regresión logística.

De esta regresión se obtiene un modelo predictivo que se utiliza para evaluar la probabilidad de que un cliente sin operaciones en default entre en default en alguna de sus operaciones de activo en el plazo de 12 meses.

El modelo utiliza un conjunto de características de predicción para predecir la probabilidad. Los clientes se clasifican como bueno o malo y se les asigna un valor de "0" a los casos que no entran en default ("bueno") o "1" a los casos que sí entran en default en la ventana de tiempo de 12 meses de observación del evento a predecir ("malo"). En este algoritmo se predice la probabilidad de que el cliente sea "malo" ( $Y=1$ ).

La regresión logística es un modelo lineal generalizado. Sirve para transformar la probabilidad, con valores comprendidos entre  $[0,1]$  en un rango mucho más amplio  $(-\infty, +\infty)$ , que hace que el valor transformado más adecuado se pueda predecir con una función lineal. La ecuación para la transformación logística de una probabilidad de que un cliente sea malo es la siguiente:

$$\text{Logit} [P(Y = 1)] = \ln \left[ \frac{P(Y = 1)}{1 - P(Y = 1)} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k = \beta_0 + \sum_{i=1}^k \beta_i x_i$$

Dónde,

$\ln$  = logaritmo natural

$P$  = Probabilidad de un cliente "malo" ( $Y = 1$ ), según variables explicativas



$1 - P$  = Probabilidad de un cliente "bueno" ( $Y = 0$ ), según variables explicativas  
 $x$  = Variables explicativas  
 $\beta_0$  = Constante de la regresión lineal  
 $\beta_k$  = Parámetros

El valor de  $P$  se puede calcular a partir de la ecuación de la regresión. La fórmula a aplicar es la siguiente:

$$P = \frac{\exp(\beta_0 + \sum_{i=1}^k \beta_i x_i)}{1 + \exp(\beta_0 + \sum_{i=1}^k \beta_i x_i)} = \frac{e^{\beta_0 + \sum_{i=1}^k \beta_i x_i}}{1 + e^{\beta_0 + \sum_{i=1}^k \beta_i x_i}}$$

Los parámetros  $\beta_1$  a  $\beta_k$  son estimados a partir de la máxima verosimilitud. Debido a las complicadas traducciones algebraicas los coeficientes no son fáciles de interpretar. En la regresión logística, los odds ratio expresados como  $\exp(\beta)$  se utilizan para medir la variación de la probabilidad de que un cliente sea "malo" por un cambio unitario en la variable predictiva (considerando el resto de variables constante).

## 2.4.1 Selección del modelo estadístico

En los siguientes apartados se describe el proceso seguido en el desarrollo del modelo estadístico inicial.

La estimación de los parámetros del modelo se ha realizado mediante el método de estimación de máxima verosimilitud.

### 2.4.1.1 Reducción de la dimensión

Una vez seleccionadas las variables candidatas para predecir la morosidad se incorporan todas en la regresión.

El criterio de selección de variables escogido para esta fase ha sido el procedimiento *Stepwise* que permite evaluar la inclusión de todas las variables candidatas una a una tomando aquellas que mejor discriminen.

Este procedimiento necesita de la especificación de los siguientes parámetros:

- **Probabilidad de Entrada (PIN= 0.05 por defecto):** Parámetro que especifica la mínima probabilidad del estadístico score para que una variable entre a formar parte del modelo. Una variable podrá entrar en el modelo si su significación es inferior a PIN.
- **Probabilidad de Salida (POUT=0.10 por defecto):** Parámetro que especifica la máxima probabilidad del estadístico condicional para que una variable que pertenece al modelo pueda ser eliminada del mismo. La variable podrá ser eliminada si su significación es superior a POUT.

El método *Stepwise* funciona siguiendo los siguientes pasos:

- Todas las variables especificadas en la regresión logística son testadas una a una mediante el estadístico score. La variable con la menor significación (y que

esta sea menor a la significación prefijada en PIN) es la primera que entra en el modelo. La distribución asintótica que sigue el estadístico score es una Chi-cuadrado con tantos grados de libertad como categorías tenga en caso de ser categórica.

$\gamma$ : Conjunto de variables que no están en el modelo.

$\alpha$ : Conjunto de variables que están en el modelo.

Si la variable es continua:  $S_i = (L_{\gamma_i}^*)^2 B_{22,i}$

Si la variable es categórica:  $S_i = (L_{\gamma_i}^*)' B_{22,i} (L_{\gamma_i}^*)$

Siendo:  $B_{22,i} = \gamma_i' \hat{V} \gamma_i - \gamma_i' \hat{V} \alpha' (\alpha' \hat{V} \alpha)^{-1} \alpha' \hat{V} \alpha$

$$\hat{V} = \text{Diag}(\hat{\pi}_1(1 - \hat{\pi}_1), \dots, \hat{\pi}_n(1 - \hat{\pi}_n))$$

L: verosimilitud de la variable  $\gamma_i$

- Después de cada entrada, las variables que están en la regresión son testadas mediante el Estadístico Condicional. La variable con una significación mayor al estadístico prefijado en POUT se elimina de la regresión y el modelo se reestima de nuevo. Las variables en el modelo son evaluadas otra vez, cuando no hay más variables que cumplan el criterio de salida, las covariables que no están en el modelo son evaluadas para entrar. El modelo termina cuando no hay más variables para entrar o salir.

Estadístico condicional:

$$-2(L(\tilde{\beta}_{(i)}) - L(full)) \sim \chi_q^2$$

Siendo: q el número de diferencia de parámetros.

$L(\tilde{\beta}_{(i)})$ : log verosimilitud evaluada con  $\tilde{\beta}_{(i)}$ .

Con este método se obtienen los parámetros del modelo estadísticamente significativos y no altamente correlacionados.

### 2.4.1.2 Significación de los parámetros del modelo

La significación de las variables del modelo se ha medido mediante el estadístico de Wald. El test que se plantea es el siguiente:

$$H_0 : \hat{\beta}_i = 0$$

$$H_1 : \hat{\beta}_i \neq 0$$

El estadístico de Wald cuando la variable no es categórica, se define del siguiente modo:

$$Wald_i = \frac{\hat{\beta}_i^2}{\hat{\sigma}_{\hat{\beta}_i}^2} \sim \chi_1^2 \quad Wald_i = \frac{\hat{\beta}_i}{\hat{\sigma}_{\hat{\beta}_i}} \sim N(0,1)$$

Siendo:

$\hat{\beta}_i$  la estimación máximo verosímil del parámetro  $\beta_i$ .

$\hat{\sigma}_{\hat{\beta}_i}^2$  varianza asintótica de  $\hat{\beta}_i$ .

El estadístico de Wald cuando la variable es categórica ( $m$  categorías), se define del siguiente modo:

$$Wald_i = \hat{\beta}_i^T V(\hat{\beta}) \hat{\beta}_i \sim \chi_m^2$$

Siendo:

$\hat{\beta}_i$  la estimación máximo verosímil del vector de dimensión  $m-1$  de parámetros de  $\beta_i$ .

$V(\hat{\beta})$ : matriz de covarianzas asintótica del vector de parámetros  $\hat{\beta}_i$ .

Ambos estadísticos se distribuyen asintóticamente como una distribución Chi-cuadrado con tantos grados de libertad como el número de parámetros estimado.

### 2.4.1.3 Análisis de multicolinealidad de las variables del modelo

Para garantizar la independencia multivariante es necesario analizar que no haya multicolinealidad en las variables que componen el modelo.

Para ello, AIS analiza el factor de inflación de la varianza (VIF) para las variables:

$$VIF_j = \frac{1}{\text{tolerancia}_j} = \frac{1}{1 - R_j^2}$$

Siendo  $R^2$  el coeficiente de determinación de la regresión de la variable “j” como variable a explicar y el resto de variables como variables explicativas.

	tolerancia	VIF
Baja multicolinealidad	$\leq 0.10$	$\leq 5$
Moderada multicolinealidad	$> 0.10 \leq 0.20$	$> 5 \leq 10$
Elevada multicolinealidad	$> 0.20$	$> 10$

De forma general, no se admiten modelos con VIF superior a 10 y se revisan aquellos con variables de multicolinealidad moderada.

### 2.4.1.4 Significación global del modelo

Se mide la significación global del modelo de  $k$  parámetros, mediante el estadístico “Likelihood ratio test”:

$$2(L_A - L_0) = 2 \ln \left( \frac{l_A}{l_0} \right) \sim \chi_k^2$$

$L_A$ : log verosimilitud del modelo de  $k$  parámetros.

$L_0$ : log verosimilitud del modelo nulo (sólo con la constante).

$l_A$ : verosimilitud del modelo de  $k$  parámetros.

$l_0$ : verosimilitud del modelo nulo (sólo con la constante).

El test planteado para un modelo de  $k$  parámetros es el siguiente:

$$H_0 : y = \hat{\beta}_0 + e$$

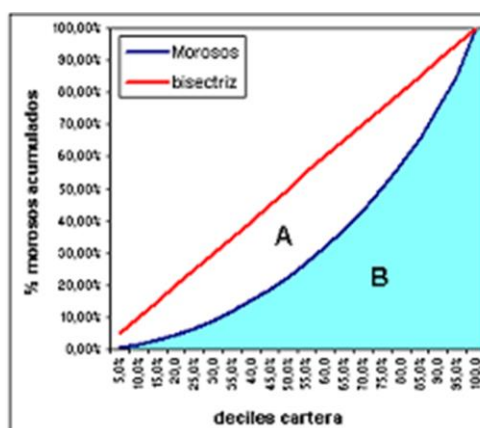
$$H_1 : y = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \dots + \hat{\beta}_k x_k + e$$

### 2.4.1.5 Poder discriminante del modelo

El poder discriminante del modelo se ha medido mediante el **coeficiente de Gini**. Este coeficiente mide el grado de concentración de un atributo en una población, en este caso estamos midiendo el grado de concentración del scoring de las operaciones malas en la cartera. Si este coeficiente tomase valor cero, es decir cuando la curva coincidiese con la bisectriz, estaríamos delante un caso de no concentración, es decir el scoring de las operaciones malas se distribuirían uniformemente y no podríamos detectar diferencias con el scoring de las operaciones buenas, ante esta situación el scoring no distinguiría entre malos y buenos.

Gráficamente, el coeficiente de GINI es el cociente entre 2 áreas:

$$\text{GINI} = \frac{A}{A + B}$$



Según la experiencia de AIS, se pueden clasificar los modelos según el coeficiente de Gini obtenido según la siguiente escala, aplicable sólo para modelos de tipo comportamental (incluyen comportamiento del cliente):

	Rango	
	Mínimo	Máximo
<b>Algoritmo de limitada capacidad de discriminación</b>		25%
<b>Algoritmo de capacidad de discriminación regular</b>	+25%	40%
<b>Algoritmo de capacidad de discriminación buena</b>	+40%	60%
<b>Algoritmo de capacidad de discriminación muy buena</b>	+60%	80%
<b>Algoritmo de capacidad de discriminación excelente</b>	+80%	

#### 2.4.1.6 Opción OFFSET

Cuando una variable tiene un poder de predicción muy elevado, esto se suele traducir en el hecho de que dicha variable tiene una contribución muy elevada dentro del modelo. El hecho de que una variable tenga demasiado peso puede ser peligroso por el hecho de que si esta variable dejara de “funcionar” en algún momento una vez implementado el modelo, la puntuación del modelo se podría ver muy castigada por este hecho y podría llevar a cálculos de puntuación incoherentes. Además, el hecho de que una variable tenga un poder de predicción muy elevado, es decir que explique muy bien la variable respuesta del modelo, también puede implicar que su presencia en el modelo base impida la presencia de otras variables con menos poder de predicción pero que podrían explicar una parte de lo que explica esta variable. Con el objetivo de mitigar la contribución de variables con un peso elevado dentro del modelo, se utiliza la opción OFFSET del paquete estadístico SAS. Esta opción funciona de la siguiente manera:

- Una vez detectada la variable o variables cuyo peso es elevado dentro del modelo. Estas se sacan del modelo y se lanza la regresión logística de nuevo, obteniendo así un nuevo modelo que no incluye dichas variables. De esta forma, las variables que se encuentran ahora en el modelo, están siendo explotadas al máximo, es decir, todo lo que quedaría explicado mediante la variable o variables con contribución elevada y que también puede ser explicado por las otras variables del modelo, ahora lo explican estas variables.
- Cuando se ha obtenido este primer modelo, la puntuación obtenida a partir de éste se fija, de tal manera que el hecho de añadir las variables que se han sacado previamente no hará variar las betas del primer modelo obtenido. Mediante la opción OFFSET, se hace otro modelo añadiendo la variable o variables que faltan, de manera que se obtiene una constante y un coeficiente asociado a dichas variables (la puntuación del modelo anterior queda fijada con coeficiente = 1).

## 2.4.2 Validación del algoritmo

### 2.4.2.1 Out of Sample

Según esta técnica de validación, se reserva un porcentaje de muestra para la validación del modelo. Así, típicamente seleccionamos un 80% aleatorio de los casos para el desarrollo del modelo y comprobamos que los resultados obtenidos son estables en la muestra del 20% restante. Así, tanto la muestra de aprendizaje como la de validación se obtienen de un mismo universo y periodo de tiempo.

Para ver si un modelo tiene un poder predictivo significativamente diferente en muestras distintas, en primer lugar se calcula el Gini para ambas muestras y su error estándar. Dada la propiedad siguiente del error estándar  $SE(Gini_{80} - Gini_{20}) = \sqrt{SE(Gini_{80})^2 + SE(Gini_{20})^2}$ , el intervalo de confianza para la diferencia de Gini al nivel  $1-\alpha$  es

$$(Gini_{80} - Gini_{20}) \pm Z_{\alpha} SE(Gini_{80} - Gini_{20}), \text{ con probabilidad } (1 - \alpha) \text{ y } Z_{\alpha} \text{ tal que } P(|N(0,1)| > Z_{\alpha}) = \alpha$$

En nuestro caso consideraremos  $1 - \alpha = 0,95$  y  $Z_{\alpha} = 1,96$ . Entonces

$$\text{Intervalo\_confianza} = (Gini_{80} - Gini_{20}) \pm 1,96 \sqrt{SE(Gini_{80})^2 + SE(Gini_{20})^2}$$

Si el intervalo anterior **no contiene el cero**, afirmamos que hay una diferencia significativa del poder de predicción del modelo a un nivel de confianza de  $1-\alpha$ .

Otras técnicas de validación con otras muestras son:

- **Out of sample – Out of time:** Ambas muestras provienen de un mismo universo pero de periodos de tiempo distintos.
- **Out of sample – Out of universe:** Ambas muestras provienen de un mismo periodo de tiempo pero universos distintos (por ejemplo de distintos segmentos de población).
- **Out of sample – Out of time – Out of universe:** Las muestras provienen de distintos universos y periodos de tiempo.

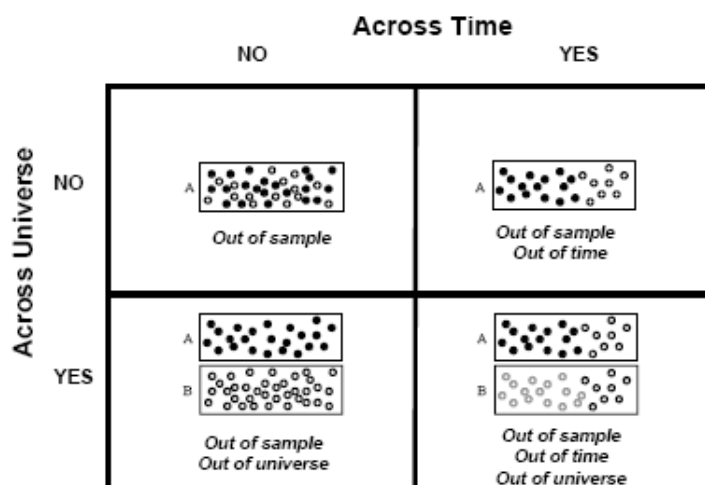


Figure 4 Schematic of out of sample validation techniques

Testing strategies are broken out based on whether they account for variances across time (horizontal axis) and across the data universe (vertical axis). Dark circles represent training data and white circles represent testing data. Gray circles represent data that may or may not be used for testing. (Reproduced from (Dhar and Stein, 1998).)

## 3 Descripción del modelo

A continuación se describe el algoritmo Proactivo obtenido a través de cada una de las variables que forman parte del mismo así como el algoritmo de integración del Modelo Proactivo con los grupos de riesgo que se obtienen de la puntuación del Modelo de Buró.

### 3.1 Modelo Proactivo

El término constante del modelo Proactivo es la siguiente:

Constante	-1.90648980
-----------	-------------

#### 3.1.1 Información de Cabecera

##### 3.1.1.1 W4\_edad- Edad (Años)

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Edad	WOE	Coeficiente	Puntuación_Edad
Más de 49	0.397555	-0.71322878	-0.28354767
De 42+ a 49	0.133797		-0.09542787
De 30+ a 42	-0.08179		0.05833498
De 25+ a 30	-0.3702		0.26403729
Hasta 25	-0.5981		0.42658213

#### Casuísticas Especiales

Si la variable viene sin informar, entonces se asignará la peor puntuación para esta variable, es decir: 0.42658213 puntos.

##### 3.1.1.2 W4\_antig\_cliente\_2- Antigüedad del cliente (Meses)

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Antigüedad	WOE	Coeficiente	Puntuación_AntigCliente
Más de 202	0.873599	-0.34133248	-0.29818771
De 77+ a 202	0.299247		-0.10214272
De 63+ a 77	0.083687		-0.02856509
De 26+ a 63	-0.20465		0.06985369
De 9+ a 26	-0.39285		0.13409246
Hasta 9	-0.55222		0.18849062



## Casuísticas Especiales

Si la variable viene sin informar, entonces se asignará la peor puntuación para esta variable, es decir: 0.18849062 puntos.

### 3.1.2 Información de Banca por Internet

#### 3.1.2.1 W4\_med\_opmon- Media número de operaciones monetarias de los últimos 6 meses

Se han utilizado las siguientes variables del bloque **Banca por Internet**, entregadas por BBVA Continental:

- OPMON1: NUMERO OPERACIONES MONETARIAS 1 MES ANTES
- OPMON2: NUMERO OPERACIONES MONETARIAS 2 MESES ANTES
- OPMON3: NUMERO OPERACIONES MONETARIAS 3 MESES ANTES
- OPMON4: NUMERO OPERACIONES MONETARIAS 4 MESES ANTES
- OPMON5: NUMERO OPERACIONES MONETARIAS 5 MESES ANTES
- OPMON6: NUMERO OPERACIONES MONETARIAS 6 MESES ANTES

Con estas variables se calcula la media del número de operaciones monetarias de los últimos 6 meses (dividiendo por el número de meses informados).

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

# Operac. Mon	WOE	Coeficiente	Puntuación_MedOpmon
Missing	-0.05536	-0.67700732	0.03747913
Más de 10	2.172087		-1.47051880
Hasta 10	0.833218		-0.56409469

### 3.1.3 Información de Pasivos

#### 3.1.3.1 W4\_med\_inv- Media del saldo medio inversión de los últimos 6 meses (soles)

Se define **Saldo medio inversión** de un mes concreto a partir de la suma de las siguientes variables entregadas por BBVA Continental (Ejemplo: 1 mes antes):

- BOLS01: SALDO MEDIO PLAZO 1 MESES ANTES
- FFFM01: SALDO MEDIO FONDOS MUTUOS 1 MESES ANTES
- PLAZ01: SALDO MEDIO PLAZO 1 MESES ANTES

Una vez obtenido el Saldo medio inversión para cada uno de los últimos 6 meses, finalmente se calcula la media (dividiendo por el número de meses informados).

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Saldo medio Invers.	WOE	Coeficiente	Puntuación_MedInv
Más de 32,431.75	2.041306	-0.38452627	-0.78493578
De 5,185+ a 32,431.75	1.333109		-0.51261543
Hasta 5,185	0.599997		-0.23071461
Missing	-0.09901		0.03807195

### 3.1.3.2 W4\_med\_ccte- Media del saldo medio cuenta corriente de los últimos 6 meses (soles)

Se han utilizado las siguientes variables entregadas por BBVA Continental:

- CCTE01: SALDO MEDIO CTA CTE 1 MESES ANTES
- CCTE02: SALDO MEDIO CTA CTE 2 MESES ANTES
- CCTE03: SALDO MEDIO CTA CTE 3 MESES ANTES
- CCTE04: SALDO MEDIO CTA CTE 4 MESES ANTES
- CCTE05: SALDO MEDIO CTA CTE 5 MESES ANTES
- CCTE06: SALDO MEDIO CTA CTE 6 MESES ANTES

Con estas variables se calcula la media del saldo media en cuenta corriente de los últimos 6 meses (dividiendo por el número de meses informados).

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Saldo Medio CCTE	WOE	Coeficiente	Puntuación_MedCCTE
Missing	-0.04217	-0.33567131	0.01415526
Más de 2,354	2.178122		-0.73113307
De 473+ a 2,354	1.1676		-0.39192982
Hasta 473	0.503508		-0.16901319

### 3.1.3.3 W4\_ahordeu\_v3- Porcentaje media saldo medio ahorros de los últimos 6 meses sobre deuda en BBVA del último mes disponible

Se han utilizado las siguientes variables entregadas por BBVA Continental:

- AHOR01: SALDO MEDIO AHORROS 1 MESES ANTES
- AHOR02: SALDO MEDIO AHORROS 2 MESES ANTES
- AHOR03: SALDO MEDIO AHORROS 3 MESES ANTES
- AHOR04: SALDO MEDIO AHORROS 4 MESES ANTES
- AHOR05: SALDO MEDIO AHORROS 5 MESES ANTES
- AHOR06: SALDO MEDIO AHORROS 6 MESES ANTES
- SDO\_BBVA\_00: DEUDA BBVA ÚLTIMO MES DISPONIBLE

Con las variables de Ahorros se calcula la media del saldo medio ahorros de los últimos 6 meses (dividiendo por el número de meses informados).

Finalmente se divide la media del saldo medio ahorros por la deuda en BBVA del último mes disponible y se multiplica por 100 para expresarlo en porcentaje. Es decir:

$$\frac{MEAN(AHOR01, AHOR02, AHOR03, AHOR04, AHOR04, AHOR06)}{SDO\_BBVA\_00} \cdot 100$$

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

% Media Ahorros sobre Deuda	WOE	Coeficiente	Puntuación_Ahordeu
No posee deuda BBVA en el último mes disponible (RCC)	0.587773	-0.38384281	-0.22561244
No posee ahorros	-0.26661		0.10233633
Más de 408,7814%	1.500944		-0.57612656
De 86,03352%+ a 408,7814%	0.754859		-0.28974720
De 20,16486%+ a 86,03352%	0.373165		-0.14323670
Hasta 20,16486%	-0.00125		0.00047980

### 3.1.4 Información de Tarjeta de Crédito

#### 3.1.4.1 w4\_med\_muso\_ctc- Media monto medio por uso compras TC mensual de los últimos 6 meses (soles)

Se han utilizado las siguientes variables entregadas por BBVA Continental para calcular el monto medio por uso en compras TC de cada uno de los 6 meses (Ejemplo: 1 mes antes):

- SC01: MONTO EN COMPRAS TC 1 MES ANTES
- UC01: USO COMPRAS TC 1 MES ANTES

El monto medio por uso en compras 1 mes antes se calcula de la siguiente manera:

$$\frac{SC01}{UC01}$$

Si el cliente posee TC y no la ha usado en compras, entonces el monto medio por uso en compras toma valor 0.

Se hace el mismo cálculo para cada uno de los seis meses y finalmente se calcula la media (dividiendo por el número de meses informados).

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Monto medio compras TC	WOE	Coeficiente	Puntuación_MedMusoCTC
Missing	0.637324	-0.46728069	-0.29780920
Más de 129,6788	0.392064		-0.18320394
De 69,56944+ a 129,6788	0.084273		-0.03937915
De 36,26389+ a 69,56944	-0.20401		0.09532993
Hasta 36,26389	-0.43444		0.20300542

### 3.1.4.2 w4\_max\_ue\_2\_tc- Máximo uso disposición efectivo TC mensual de los últimos 6 (usos)

Se han utilizado las siguientes variables entregadas por BBVA Continental:

- UE01: USO DISPOSICION EFECTIVO 1 MES ANTES
- UE02: USO DISPOSICION EFECTIVO 2 MESES ANTES
- UE03: USO DISPOSICION EFECTIVO 3 MESES ANTES
- UE04: USO DISPOSICION EFECTIVO 4 MESES ANTES
- UE05: USO DISPOSICION EFECTIVO 5 MESES ANTES
- UE06: USO DISPOSICION EFECTIVO 6 MESES ANTES

A partir de estas variables se calcula el máximo uso en disposición de efectivo de los últimos 6 meses.

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Máx. Uso Disp. Efect. TC	WOE	Coeficiente	Puntuación_MaxUeTC
Missing	0.658301	-0.38370642	-0.25259432
0	0.316816		-0.12156433
1	-0.20736		0.07956536
De 1+ a 3	-0.4452		0.17082610
De 3+ a 5	-0.77741		0.29829721
Más de 5	-0.97353		0.37354971

## 3.1.5 Información de Préstamos

### 3.1.5.1 w4\_riesgo\_pasv4- Porcentaje cuotas préstamos más media pago TC últimos 6 meses sobre media saldo medio pasivo vista + CTS

Se define **Saldo medio pasivo vista + CTS** de un mes concreto a partir de la suma de las siguientes variables entregadas por BBVA Continental (Ejemplo: 1 mes antes):

- AHOR01: SALDO MEDIO AHORROS 1 MESES ANTES
- CCTE01: SALDO MEDIO CTA CTE 1 MESES ANTES
- CTS01: SALDO MEDIO CTS 1 MES ANTES

Una vez obtenido el Saldo medio pasivo vista + CTS para cada uno de los últimos 6 meses, finalmente se calcula la media (dividiendo por el número de meses informados).

La media del pago TC se calcula a partir de las siguientes variables:

- PT01
- PT02
- PT03
- PT04
- PT05
- PT06

Con estas variables se calcula la media de pago TC de los últimos 6 meses (dividiendo por el número de meses informados).

Para el cálculo de la cuota total de préstamos se utilizan las siguientes variables:

- HIPCUO: CUOTA MEDIA PRÉSTAMOS HIPOTECARIOS
- HIPCON: NÚMERO DE PRÉSTAMOS HIPOTECARIOS
- VEHCUO: CUOTA MEDIA PRÉSTAMOS VEHICULARES
- VEHCON: NÚMERO DE PRÉSTAMOS VEHICULARES
- PERCUO: CUOTA MEDIA PRÉSTAMOS PERSONALES
- PERCON: NÚMERO DE PRÉSTAMOS PERSONALES

Así, Cuota Total Préstamos =  $HIPCUO \cdot HIPCON + VEHCUO \cdot VEHCON + PERCUO \cdot PERCON$

Finalmente, la variable se define de la siguiente manera:

$$\frac{Cuota\ Total\ Préstamos + Media\ Pago\ TC}{Media\ Pasivo\ Vista + CTS} \cdot 100$$

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

% cuotas mensuales / SMPV +CTS	WOE	Coeficiente	Puntuación_RiesgoPasV
Hasta 2.10279%	0.936058	-0.19784165	-0.18519126
De 2.10279+ a 6.18238%	0.803605		-0.15898654
De 6.18238+ a 19.09661%	0.551284		-0.10906694
De 19.09661+ a 283.2669%	0.2592		-0.05128056
Más de 283.2669%	-0.10409		0.02059334
Media Pasivo Vista + CTS = 0 ó Missing	-0.2425		0.04797660

### Casuísticas Especiales

Si el numerador es igual a 0 ó Missing, entonces la variable tomará valor 0 y se asignará la puntuación del tramo “Hasta 2.10279”, es decir -0.18519126 puntos.

### 3.1.6 Información de RCC

#### 3.1.6.1 w4\_mean\_conctjpear- Media del porcentaje de uso tarjeta Pear Group de los últimos 6 meses

Se define el porcentaje de uso de la tarjeta en el Pear Group de un mes concreto de la siguiente manera:

$$\frac{\text{Línea utilizada BBVA} + \text{IBK} + \text{BCP} + \text{SCO}}{\text{Línea concedida BBVA} + \text{IBK} + \text{BCP} + \text{SCO}} \cdot 100$$

Se calcula el porcentaje de uso de la tarjeta en el Pear Group para cada uno de los últimos 6 meses disponibles y finalmente se calcula la media (dividiendo por el número de meses informados).

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

% uso Pear Group	WOE	Coeficiente	Puntuación_MeanPear
Missing	0.579623	-0.58890011	-0.341340048
Hasta 10.56051%	0.934321		-0.55022174
De 10.56051%+ a 17.51088%	0.678557		-0.399602292
De 17.51088%+ a 33.72041%	0.387729		-0.228333651
De 33.72041%+ a 42.34977%	0.089339		-0.052611747
De 42.34977%+ a 51.47676%	-0.13987		0.082369458
De 51.47676%+ a 70.4578%	-0.41614		0.245064892
De 70.4578%+ a 80.71603%	-0.70222		0.413537435
Más de 80.71603%	-0.90041		0.530251548

#### 3.1.6.2 w4\_mean\_tjfalrip- Media porcentaje uso tarjeta Falabella + Ripley de los últimos 6 meses

Se define el porcentaje de uso de la tarjeta en Falabella y Ripley de un mes concreto de la siguiente manera:

$$\frac{\text{Línea utilizada Falabella} + \text{Ripley}}{\text{Línea concedida Falabella} + \text{Ripley}} \cdot 100$$

Se calcula el porcentaje de uso de la tarjeta en Falabella y Ripley para cada uno de los últimos 6 meses disponibles y finalmente se calcula la media (dividiendo por el número de meses informados).

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

% uso TCR+TCF	WOE	Coeficiente	Puntuación_MeanFalRip
Missing	-0.04737	-0.52917516	0.02506703
Hasta 6.745401%	1.318281		-0.69760156
De 6.745401%+ a 16.92315%	0.963194		-0.50969834
De 16.92315%+ a 25.60236%	0.590019		-0.3122234
De 25.60236%+ a 39.92466%	0.26949		-0.14260741
De 39.92466%+ a 49.89152%	0.00811		-0.00429161
De 49.89152%+ a 69.48034%	-0.21695		0.11480455
De 69.48034%+ a 84.79178%	-0.55063		0.29137972
Más de 84.79178%	-0.83123		0.43986627

### 3.1.7 Información de Comportamiento

El comportamiento de pago del cliente en la entidad en los últimos 12 meses se ha recogido mediante dos variables, una referente al último trimestre (últimos 3 meses) y otra referente a los 9 meses anteriores al último trimestre.

#### 3.1.7.1 W4\_peor\_sitBBVA3m- Peor Situación BBVA del último trimestre

Variable categórica que puede tomar los siguientes valores en función de los días de impago registrados en el BBVA en los últimos 3 meses:

- 0 = El máximo días de impago en el BBVA durante los últimos 3 meses ha sido 0
- 1 = El máximo días de impago en el BBVA durante los últimos 3 meses ha sido superior o igual a 1 día e inferior a 15 días
- 2 = El máximo días de impago en el BBVA durante los últimos 3 meses ha sido superior o igual a 15 días e inferior a 30 días
- 3 = El máximo días de impago en el BBVA durante los últimos 3 meses ha sido superior o igual a 30 días e inferior a 60 días
- 4 = El máximo días de impago en el BBVA durante los últimos 3 meses ha sido superior o igual a 60 días

El coeficiente asociado a esta variable ha sido asignado de manera experta de forma que la variable tenga una contribución de alrededor de un 15% dentro del modelo Proactivo.

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Peor Situación - 3m	WOE	Coeficiente	Puntuación_Sit3m
0	0.941196	-0.4849099	-0.456395258
1	0.022466		-0.010893986
2	-1.08652		0.526864305
3 y 4	-1.62121		0.786140779

### 3.1.7.2 W4\_peor\_sitBBVA9m- Peor situación BBVA en los 9 meses anteriores al último trimestre

Variable categórica que puede tomar los siguientes valores en función del máximo de días de impago registrados en el BBVA en los 9 meses anteriores al último trimestre:

0 = El máximo días de impago en el BBVA en los 9 meses anteriores al último trimestre ha sido 0

1 = El máximo días de impago en el BBVA en los 9 meses anteriores al último trimestre ha sido superior o igual a 1 día e inferior a 15 días

2 = El máximo días de impago en el BBVA en los 9 meses anteriores al último trimestre ha sido superior o igual a 15 días e inferior a 30 días

3 = El máximo días de impago en el BBVA en los 9 meses anteriores al último trimestre ha sido superior o igual a 30 días e inferior a 60 días

4 = El máximo días de impago en el BBVA en los 9 meses anteriores al último trimestre ha sido superior o igual a 60 días

El coeficiente asociado a esta variable ha sido asignado de manera experta de forma que la variable tenga una contribución de alrededor de un 10% dentro del modelo Proactivo.

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Peor Situación – 9m	WOE	Coeficiente	Puntuación_Sit9m
0	0.510416	-0.55873173	-0.28518561
1	0.059034		-0.03298417
2	-0.5749		0.32121487
3	-0.8484		0.47402800
4	-1.01452		0.56684451

### 3.1.8 Obtención de la puntuación del Modelo Proactivo

La puntuación correspondiente al Model Proactivo para los clientes de Activo del BBVA se calcula a partir de la siguiente ecuación:

$$\begin{aligned} \text{Puntuación\_Modelo\_Proactivo} = & -1.90648980 + \text{Puntuación\_Edad} + \\ & \text{Puntuación\_AntigCliente} + \text{Puntuación\_MedOpmon} + \text{Puntuación\_MedInv} + \\ & \text{Puntuación\_MedCCTE} + \text{Puntuación\_Ahordeu} + \text{Puntuación\_MedMusoCTC} + \\ & \text{Puntuación\_MaxUeTC} + \text{Puntuación\_RiesgoPasV} + \text{Puntuación\_MeanPear} + \\ & \text{Puntuación\_MeanFalRip} + \text{Puntuación\_Sit3m} + \text{Puntuación\_Sit9m} \end{aligned}$$

## 3.2 Integración Proactivo + Grupo Riesgo Buró

Con el objetivo de integrar el Modelo Proactivo con el Grupo de Buró (obtenido a partir de los grupos de riesgo definidos mediante el Modelo de Buró), se fija la puntuación obtenida en el Modelo Proactivo mediante la opción offset y se hace otro modelo



añadiendo la variable  $w4\_COD\_CRRIESGO$ . Con este modelo se obtiene una constante y un coeficiente asociado a dicha variable (la puntuación del Proactivo queda fijada con coeficiente = 1).

El término constante del modelo de integración con el Grupo de Buró es la siguiente:

Constante	-0.07102552
-----------	-------------

### 3.2.1 $w4\_cod\_crriesgo$ - Grupo Buró

Se ha utilizado las siguiente variable entregada por BBVA Continental:

- $COD\_CRRIESGO$ : GRUPO DE RIESGO BURO

La respuesta del modelo frente a los valores de dicha variable es la siguiente:

Grupo Buró	WOE	Coeficiente	Puntuación_CodCRRiesgo
Missing	-0.02579	-0.22703392	0.00585520
Grupo 1	2.18093		-0.49514509
Grupo 2	1.47952		-0.33590123
Grupo 3	1.032392		-0.23438800
Grupo 4	0.460112		-0.10446103
Grupo 5	-0.17624		0.04001246
Grupos 6 y 7	-0.59215		0.13443814
Grupo 8	-0.96925		0.22005263

Cabe señalar que aunque trabajamos con el segmento de Clientes de Activo de BBVA, puede ser que ellos no tengan grupo de Buró por no haber tenido activos en el Sistema Financiero entre 7 y 24 meses antes de la fecha de referencia. Es decir, por ser clientes recientes de activo. Por ello, la relevancia de la categoría “missing” en la integración.

### 3.2.2 Obtención de la puntuación final del cliente

La puntuación final correspondiente a la integración del Modelo Proactivo con el Grupo de Buró para los clientes de Activo del BBVA se calcula a partir de la siguiente ecuación:

$\text{Puntuación\_Final} = -0.07102552 + \text{Puntuación\_Modelo\_Proactivo} + \text{Puntuación\_CodCRRiesgo}$
--

Una vez obtenida la puntuación final del cliente, la probabilidad de mora se obtiene aplicando la siguiente expresión, que corresponde al modelo Logit (Regresión Logística):

$$\text{Probabilidad\_de\_Mora}_i = \frac{1}{1 + e^{-Z_i}} = \frac{1}{1 + e^{-(\text{Puntuación\_Final})}}$$

### 3.3 Aspectos cualitativos del Modelo Proactivo

---

El modelo Proactivo se ha realizado en dos fases con el objetivo principal de reducir el peso total de las variables de Comportamiento a un 25% dentro del modelo a petición de BBVA-Continental. De esta forma, se han realizado los siguientes pasos:

- 1- Se ha hecho un primer modelo incluyendo todas las variables exceptuando las variables de comportamiento: *Edad (años), Antigüedad del Cliente (meses), Media número de operaciones monetarias de los últimos 6 meses, Media del saldo medio pasivo no vista de los últimos 6 meses, Media del saldo medio cuenta corriente de los últimos 6 meses, Porcentaje media saldo medio ahorros de los últimos 6 meses sobre deuda bbva del último mes disponible, Máximo uso disposición efectivo TC mensual de los últimos 6 meses, Media monto medio por uso compras TC mensual de los últimos 6 meses, Porcentaje cuotas préstamos más media pago TC últimos 6 meses sobre media saldo medi pasivo vista, Media del porcentaje uso tarjeta Pear Group de los últimos 6 meses, Media porcentaje uso tarjeta Falabella + Ripay de los últimos 6 meses.*
- 2- Se fijan los coeficientes del modelado obtenido y se asignan pesos de forma experta a las variables de comportamiento: *Peor situación BBVA del último trimestre, Peor Situación BBVA en los 9 meses anteriores al último trimestre* para que el peso de estas variables dentro del modelo sea de un 15% y un 10% respectivamente.

Los pesos han sido asignados de forma experta a petición de BBVA-Continental.

## 4 Recuento de Resultados

### 4.1 Ranking de las variables del modelo

#### 4.1.1 Ranking de las variables del modelo Proactivo

En la siguiente tabla se describe la contribución empírica (según la desviación estándar) de cada una de las variables explicativas del modelo. Se ha ordenado la tabla de mayor a menor valor de contribución empírica.

Variable	Desviación Estándar	Contribución Empírica
w4_peor_sitbbva3m	0.42	15%
w4_mean_conctjpear	0.37	13.25%
w4_peor_sitbbva9m	0.28	10%
w4_mean_tjfalrip	0.28	9.92%
w4_edad_2	0.21	7.63%
w4_med_muso_ctc	0.2	7.34%
w4_med_opmon	0.2	7.03%
w4_max_ue_2_tc	0.2	7.02%
w4_ahordeu_v3	0.16	5.81%
w4_med_pasnv	0.15	5.53%
w4_antig_cliente_2	0.13	4.73%
w4_med_ccte	0.11	3.94%
w4_riesgo_pasv4	0.08	2.82%

## 4.1.2 Ranking de las variables del modelo integración Proactivo + Grupo Riesgo Buró

En la siguiente tabla se describe la contribución empírica (según la desviación estándar) de cada una de las variables explicativas del modelo.

Se ha ordenado la tabla de mayor a menor valor de contribución empírica.

Variable	Desviación Estándar	Contribución Empírica
w4_peor_sitBBVA3m	0.42	14.05%
w4_mean_conctjpear	0.37	12.40%
w4_peor_sitBBVA9m	0.28	9.36%
w4_mean_tjfalrip	0.28	9.28%
w4_edad_2	0.21	7.14%
w4_med_muso_ctc	0.20	6.88%
w4_med_opmon	0.20	6.58%
w4_max_ue_2_tc	0.20	6.58%
w4_cod_crriesgo	0.19	6.36%
w4_ahordeu_v3	0.16	5.44%
w4_med_pasnv	0.15	5.17%
w4_antig_cliente_2	0.13	4.43%
w4_med_ccte	0.11	3.69%
w4_riesgo_pasv4	0.08	2.64%

## 4.2 Eficiencia

En este apartado se muestra la eficiencia del modelo en las siguientes muestras:

- 1) **Muestra de desarrollo del modelo**: 163,082 clientes. 21,210 morosos.
- 2) **Muestra de validación**: 40,495 clientes. 5,344 morosos.

### 4.2.1 Modelo Proactivo

#### 4.2.1.1 Muestra de desarrollo del modelo

A continuación se observan los resultados del modelo Proactivo sobre la muestra de entrenamiento:

% aprobación	Tramo inferior	Tramo superior	Buenas	Morosas	Total	% mora del grupo	% mora asumida	%buenas acum.	% de mora acum.
5.0%	0.00	0.01	8125	29	8154	0.4%	0.4%	5.7%	0.1%
10.0%	0.01	0.02	8054	89	8143	1.1%	0.7%	11.4%	0.6%
15.0%	0.02	0.02	8046	111	8157	1.4%	0.9%	17.1%	1.1%
20.0%	0.02	0.03	8025	139	8164	1.7%	1.1%	22.7%	1.7%
25.0%	0.03	0.03	7911	205	8116	2.5%	1.4%	28.3%	2.7%
30.0%	0.03	0.04	7925	265	8190	3.2%	1.7%	33.9%	4.0%
35.0%	0.04	0.05	7812	339	8151	4.2%	2.1%	39.4%	5.5%
40.0%	0.05	0.05	7723	435	8158	5.3%	2.5%	44.8%	7.6%
45.0%	0.05	0.06	7662	492	8154	6.0%	2.9%	50.2%	9.9%
50.0%	0.06	0.08	7511	643	8154	7.9%	3.4%	55.5%	13.0%
55.0%	0.08	0.09	7433	721	8154	8.8%	3.9%	60.8%	16.4%
60.0%	0.09	0.11	7291	863	8154	10.6%	4.4%	65.9%	20.4%
65.0%	0.11	0.13	7141	1013	8154	12.4%	5.0%	71.0%	25.2%
70.0%	0.13	0.15	6942	1213	8155	14.9%	5.7%	75.8%	30.9%
75.0%	0.15	0.18	6747	1407	8154	17.3%	6.5%	80.6%	37.5%
80.0%	0.18	0.22	6490	1664	8154	20.4%	7.4%	85.2%	45.4%
85.0%	0.22	0.27	6162	1992	8154	24.4%	8.4%	89.5%	54.8%
90.0%	0.27	0.34	5659	2495	8154	30.6%	9.6%	93.5%	66.5%
95.0%	0.34	0.46	4820	3335	8155	40.9%	11.3%	96.9%	82.3%
100.0%	0.46	1.00	4393	3760	8153	46.1%	13.0%	100.0%	100.0%
<b>Total</b>			<b>141872</b>	<b>21210</b>	<b>163082</b>	<b>13.0%</b>		<b>GINI =</b>	<b>60.28%</b>

#### 4.2.1.2 Muestra de validación

A continuación se observan los resultados del modelo Proactivo sobre la muestra de validación:

% aprobación	Tramo inferior	Tramo superior	Buenas	Morosas	Total	% mora del grupo	% mora asumida	% buenas acum.	% de mora acum.
5.0%	0.00	0.01	2009	15	2024	0.7%	0.7%	5.7%	0.3%
10.0%	0.01	0.02	2003	22	2025	1.1%	0.9%	11.4%	0.7%
15.0%	0.02	0.02	1993	32	2025	1.6%	1.1%	17.1%	1.3%
20.0%	0.02	0.03	1997	28	2025	1.4%	1.2%	22.8%	1.8%
25.0%	0.03	0.03	1980	60	2040	2.9%	1.5%	28.4%	2.9%
30.0%	0.03	0.04	1942	68	2010	3.4%	1.9%	33.9%	4.2%
35.0%	0.04	0.05	1944	80	2024	4.0%	2.2%	39.5%	5.7%
40.0%	0.05	0.06	1910	115	2025	5.7%	2.6%	44.9%	7.9%
45.0%	0.06	0.07	1900	125	2025	6.2%	3.0%	50.3%	10.2%
50.0%	0.07	0.08	1880	144	2024	7.1%	3.4%	55.6%	12.9%
55.0%	0.08	0.09	1852	173	2025	8.5%	3.9%	60.9%	16.1%
60.0%	0.09	0.11	1801	224	2025	11.1%	4.5%	66.0%	20.3%
65.0%	0.11	0.13	1766	259	2025	12.8%	5.1%	71.1%	25.2%
70.0%	0.13	0.16	1700	322	2022	15.9%	5.9%	75.9%	31.2%
75.0%	0.16	0.18	1671	356	2027	17.6%	6.7%	80.6%	37.9%
80.0%	0.18	0.22	1602	423	2025	20.9%	7.6%	85.2%	45.8%
85.0%	0.22	0.27	1524	501	2025	24.7%	8.6%	89.5%	55.1%
90.0%	0.27	0.35	1400	625	2025	30.9%	9.8%	93.5%	66.8%
95.0%	0.35	0.47	1166	859	2025	42.4%	11.5%	96.8%	82.9%
100.0%	0.47	1.00	1111	913	2024	45.1%	13.2%	100.0%	100.0%
<b>Total</b>			<b>35151</b>	<b>5344</b>	<b>40495</b>	<b>13.2%</b>		<b>GINI =</b>	<b>60.00%</b>

## 4.2.2 Modelo integración Proactivo + Grupo Riesgo Buró

### 4.2.2.1 Muestra de desarrollo del modelo

A continuación se observan los resultados del modelo integración Proactivo con el Grupo de Buró sobre la muestra de entrenamiento:

% aprobación	Tramo inferior	Tramo superior	Buenas	Morosas	Total	% mora del grupo	% mora asumida	% buenas acum.	% de mora acum.
5.0%	0.00	0.01	8128	26	8154	0.3%	0.3%	5.7%	0.1%
10.0%	0.01	0.01	8074	80	8154	1.0%	0.6%	11.4%	0.5%
15.0%	0.01	0.02	8067	88	8155	1.1%	0.8%	17.1%	0.9%
20.0%	0.02	0.02	8021	132	8153	1.6%	1.0%	22.8%	1.5%
25.0%	0.02	0.03	7953	201	8154	2.5%	1.3%	28.4%	2.5%
30.0%	0.03	0.03	7902	252	8154	3.1%	1.6%	33.9%	3.7%
35.0%	0.03	0.04	7818	337	8155	4.1%	2.0%	39.4%	5.3%
40.0%	0.04	0.05	7736	418	8154	5.1%	2.4%	44.9%	7.2%
45.0%	0.05	0.06	7649	505	8154	6.2%	2.8%	50.3%	9.6%
50.0%	0.06	0.07	7571	583	8154	7.1%	3.2%	55.6%	12.4%
55.0%	0.07	0.08	7424	730	8154	9.0%	3.7%	60.9%	15.8%
60.0%	0.08	0.10	7295	859	8154	10.5%	4.3%	66.0%	19.9%
65.0%	0.10	0.12	7152	1002	8154	12.3%	4.9%	71.0%	24.6%
70.0%	0.12	0.14	6977	1178	8155	14.4%	5.6%	76.0%	30.1%
75.0%	0.14	0.17	6666	1488	8154	18.2%	6.4%	80.7%	37.1%
80.0%	0.17	0.21	6543	1611	8154	19.8%	7.3%	85.3%	44.7%
85.0%	0.21	0.27	6123	2031	8154	24.9%	8.3%	89.6%	54.3%
90.0%	0.27	0.35	5633	2521	8154	30.9%	9.6%	93.6%	66.2%
95.0%	0.35	0.48	4782	3372	8154	41.4%	11.2%	96.9%	82.1%
100.0%	0.48	1.00	4358	3796	8154	46.6%	13.0%	100.0%	100.0%
<b>Total</b>			<b>141872</b>	<b>21210</b>	<b>163082</b>	<b>13.0%</b>		<b>GINI =</b>	<b>61.09%</b>

#### 4.2.2.2 Muestra de validación

A continuación se observan los resultados del modelo integración Proactivo con el Grupo de Buró sobre la muestra de validación:

% aprobación	Tramo inferior	Tramo superior	Buenas	Morosas	Total	% mora del grupo	% mora asumida	% buenas acum.	% de mora acum.
5.0%	0.00	0.01	2013	11	2024	0.5%	0.5%	5.7%	0.2%
10.0%	0.01	0.01	2007	18	2025	0.9%	0.7%	11.4%	0.5%
15.0%	0.01	0.02	1999	26	2025	1.3%	0.9%	17.1%	1.0%
20.0%	0.02	0.02	1990	34	2024	1.7%	1.1%	22.8%	1.7%
25.0%	0.02	0.03	1978	47	2025	2.3%	1.3%	28.4%	2.5%
30.0%	0.03	0.03	1954	71	2025	3.5%	1.7%	34.0%	3.9%
35.0%	0.03	0.04	1941	84	2025	4.1%	2.1%	39.5%	5.4%
40.0%	0.04	0.05	1913	112	2025	5.5%	2.5%	44.9%	7.5%
45.0%	0.05	0.06	1894	131	2025	6.5%	2.9%	50.3%	10.0%
50.0%	0.06	0.07	1885	139	2024	6.9%	3.3%	55.7%	12.6%
55.0%	0.07	0.08	1842	183	2025	9.0%	3.8%	60.9%	16.0%
60.0%	0.08	0.10	1834	191	2025	9.4%	4.3%	66.1%	19.6%
65.0%	0.10	0.12	1744	281	2025	13.9%	5.0%	71.1%	24.9%
70.0%	0.12	0.15	1718	307	2025	15.2%	5.8%	76.0%	30.6%
75.0%	0.15	0.18	1675	349	2024	17.2%	6.5%	80.8%	37.1%
80.0%	0.18	0.22	1610	415	2025	20.5%	7.4%	85.3%	44.9%
85.0%	0.22	0.27	1496	529	2025	26.1%	8.5%	89.6%	54.8%
90.0%	0.27	0.35	1393	632	2025	31.2%	9.8%	93.6%	66.6%
95.0%	0.35	0.48	1169	856	2025	42.3%	11.5%	96.9%	82.6%
100.0%	0.48	1.00	1096	928	2024	45.8%	13.2%	100.0%	100.0%
<b>Total</b>			<b>35151</b>	<b>5344</b>	<b>40495</b>	<b>13.2%</b>		<b>GINI =</b>	<b>60.77%</b>

#### 4.2.3 Colectivos de Control

Puesto que al inicio del proyecto se definieron unos colectivos estratégicos para el Banco, para los cuales se esperaba que los modelos construidos respondieran razonablemente bien, a lo largo de todo el proyecto se han ido validando los resultados de los llamados colectivos de control.

Así, se muestra en la siguiente tabla los valores del Coeficiente de Gini para dichos colectivos:

COLECTIVO	GINI	# CASOS
JOVEN	52.04%	9,136
TARJETA DE DÉBITO	67.35%	33,973
BANCA POR INTERNET	72.03%	13,181
VIP	70.51%	32,589
MUESTRA DESARROLLO	61.09%	163,082



## 5 Validación

Seguidamente se exponen los Gini obtenidos sobre la muestra de desarrollo y sobre la muestra de validación, junto con los intervalos de confianza de la diferencia de Gini que nos indica, si hay o no diferencias significativas de poder de predicción del modelo en ambas muestras.

	GINI Muestra Desarrollo	GINI Muestra Validación	Intervalo de confianza 95 %
Modelo Proactivo	60.28%	60.00%	[-0.0103,0.0159]
Modelo Final: Modelo Proactivo + Integración Buró	61.09%	60.77%	[-0.0098,0.0162]

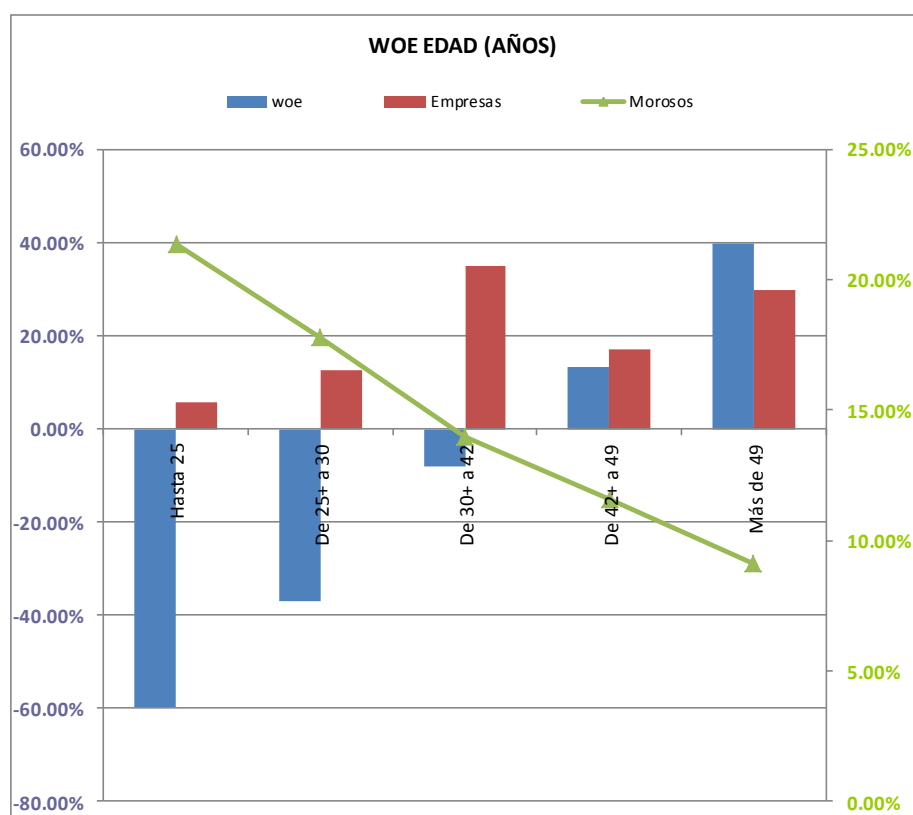
Dado que el intervalo de confianza contiene el cero, con un 95% de confianza se puede afirmar que los modelos son estables.

## 6 Anexo: gráficos woe y reporte técnico

### 6.1 Woe de las variables del modelo

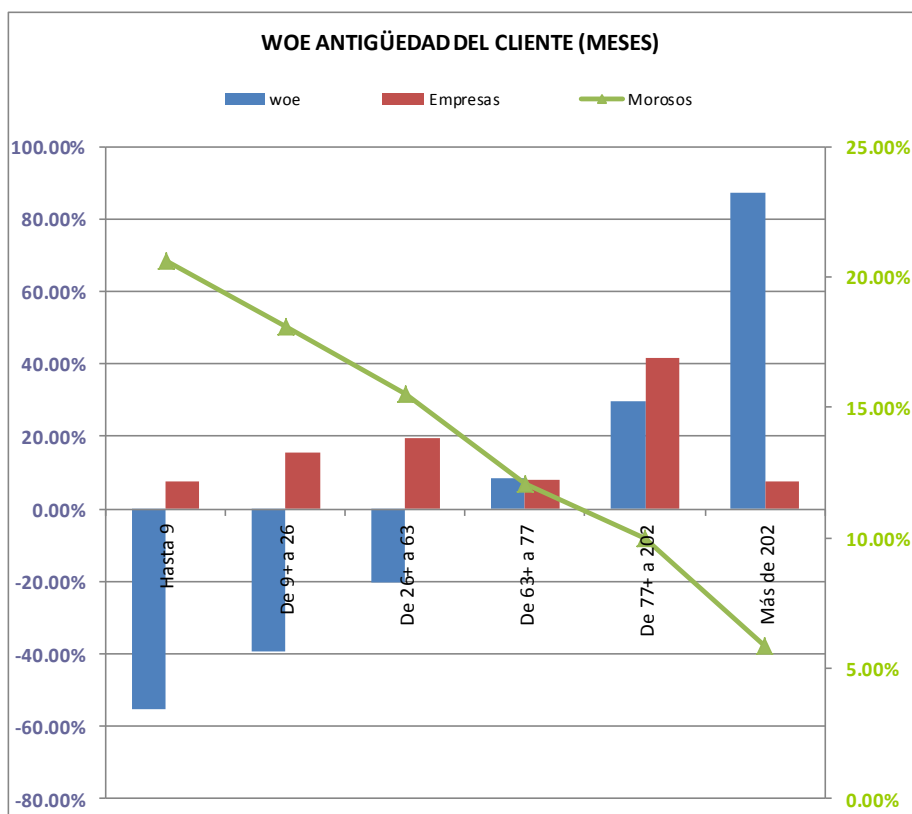
#### 6.1.1 Cabecera

##### 6.1.1.1 WOE EDAD (AÑOS)



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coficiente	Coficiente
Hasta 25	9,136	5.60%	1,953	21.38%	-59.81%	2.48%
De 25+ a 30	20,415	12.52%	3,633	17.80%	-37.02%	1.96%
De 30+ a 42	57,101	35.01%	7,971	13.96%	-8.18%	0.24%
De 42+ a 49	27,738	17.01%	3,208	11.57%	13.38%	0.29%
Más de 49	48,692	29.86%	4,445	9.13%	39.76%	4.07%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>9.04%</b>

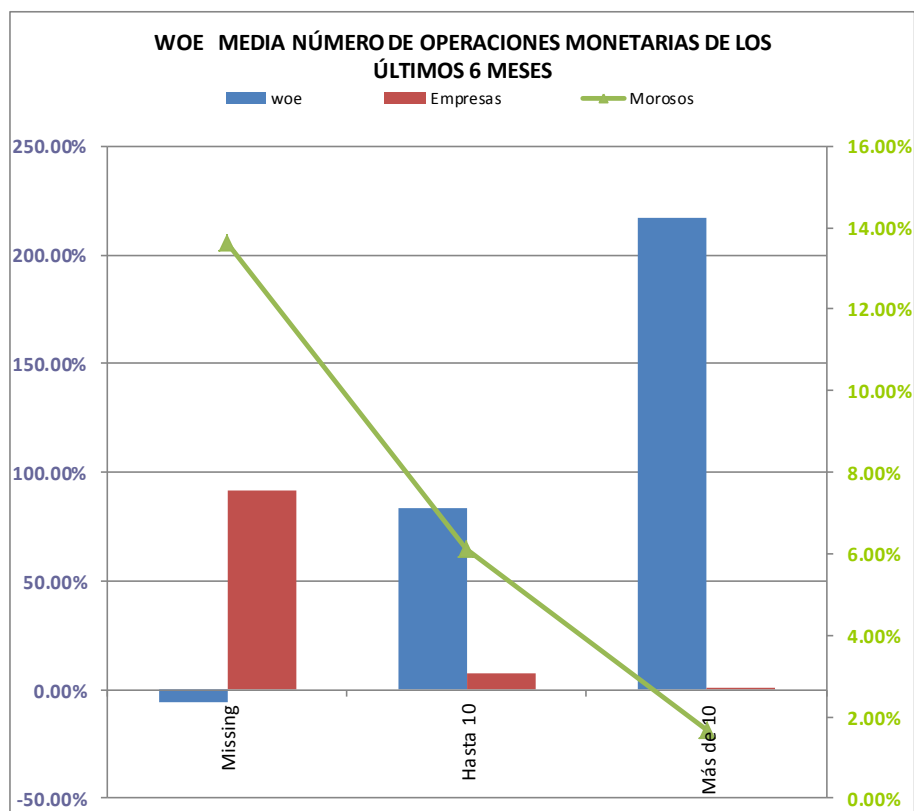
### 6.1.1.2 WOE ANTIGÜEDAD DEL CLIENTE (MESES)



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coeficiente	Coeficiente
Hasta 9	12,728	7.80%	2,624	20.62%	-55.22%	2.90%
De 9+ a 26	25,362	15.55%	4,598	18.13%	-39.29%	2.77%
De 26+ a 63	31,610	19.38%	4,900	15.50%	-20.46%	0.87%
De 63+ a 77	13,162	8.07%	1,591	12.09%	8.37%	0.05%
De 77+ a 202	67,861	41.61%	6,771	9.98%	29.92%	3.33%
Más de 202	12,359	7.58%	726	5.87%	87.36%	4.17%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	-	<b>14.10%</b>

## 6.1.2 Banca por Internet

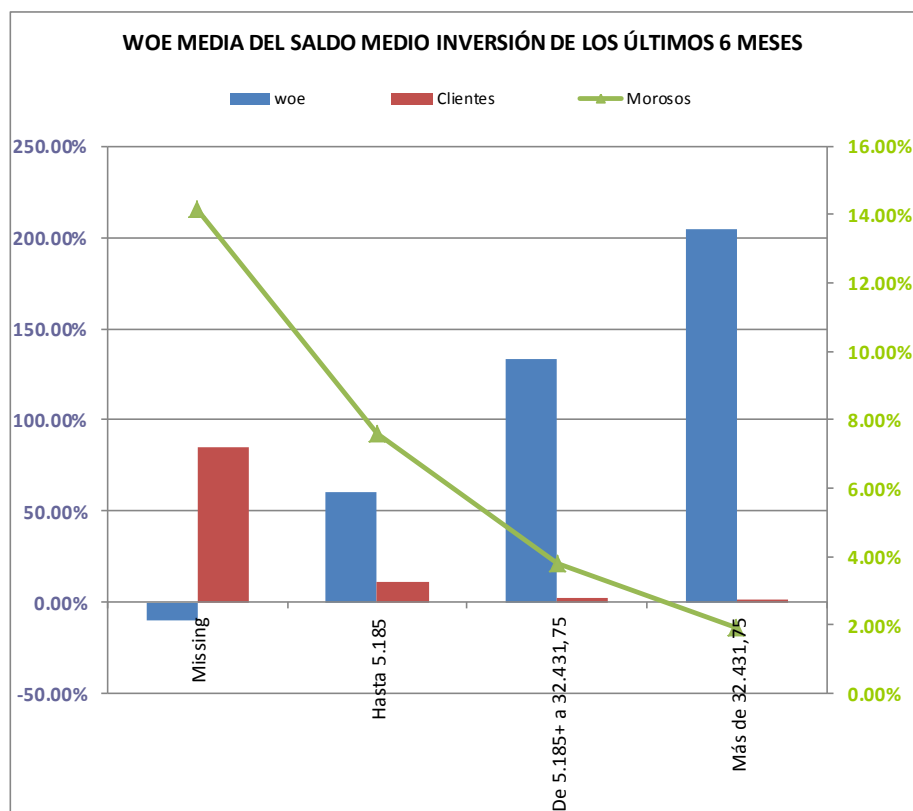
### 6.1.2.1 WOE MEDIA NÚMERO DE OPERACIONES MONETARIAS DE LOS ÚLTIMOS 6 MESES



Valor	Empresas		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coefficiente	Coefficiente
Missing	149,857	91.89%	20,448	13.65%	-5.54%	0.29%
Hasta 10	12,210	7.49%	745	6.10%	83.32%	3.81%
Más de 10	1,015	0.62%	17	1.67%	217.21%	1.35%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>5.45%</b>

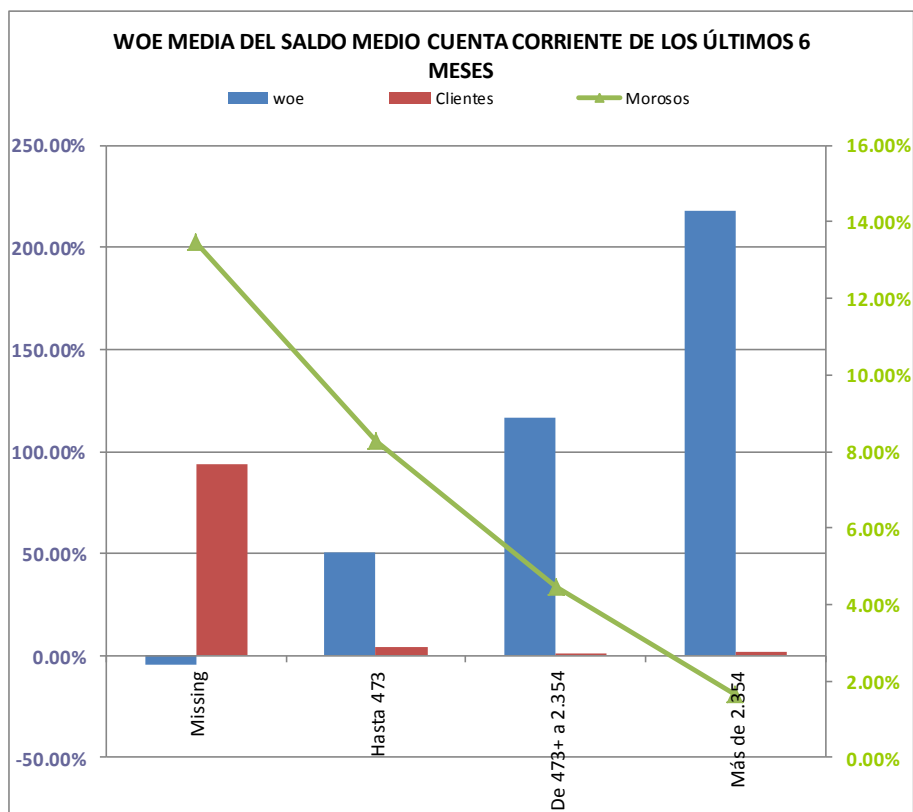
## 6.1.3 Pasivo

### 6.1.3.1 WOE MEDIA DEL SALDO MEDIO INVERSIÓN DE LOS ÚLTIMOS 6 MESES



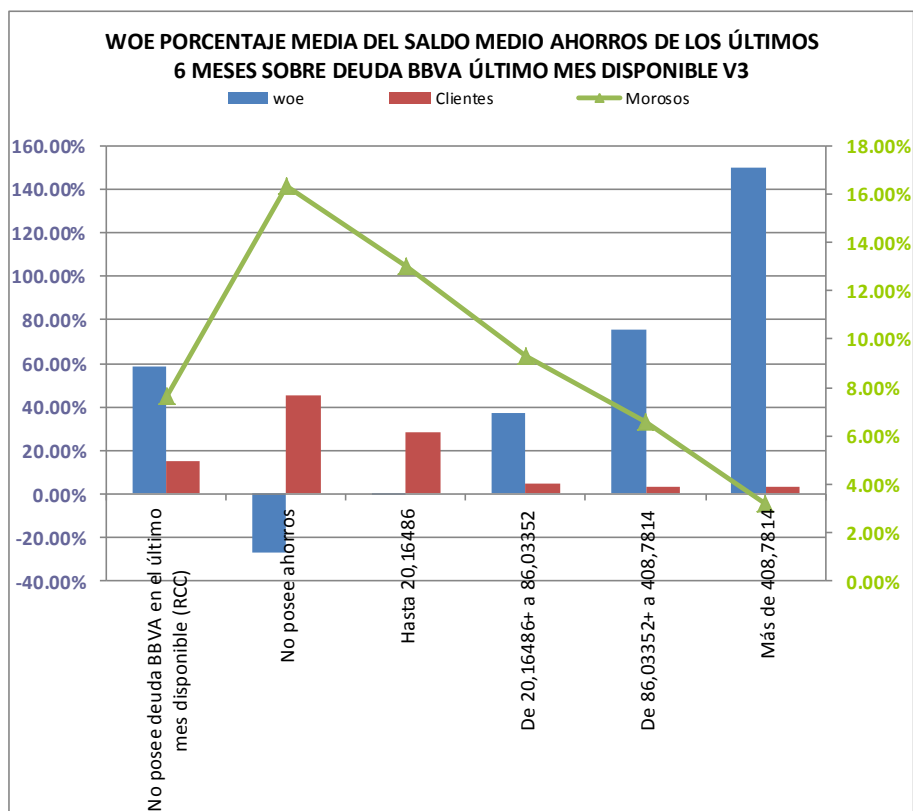
Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coficiente	Coficiente
Missing	139,015	85.24%	19,695	14.17%	-9.90%	0.87%
Hasta 5.185	17,329	10.63%	1,314	7.58%	60.00%	3.06%
De 5.185+ a 32.431,75	3,850	2.36%	146	3.79%	133.31%	2.56%
Más de 32.431,75	2,888	1.77%	55	1.90%	204.13%	3.55%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	-	<b>10.03%</b>

### 6.1.3.2 WOE MEDIA DEL SALDO MEDIO CUENTA CORRIENTE DE LOS ÚLTIMOS 6 MESES



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coefficiente	Coefficiente
Missing	152,568	93.55%	20,582	13.49%	-4.22%	0.17%
Hasta 473	6,311	3.87%	523	8.29%	50.35%	0.81%
De 473+ a 2.354	1,260	0.77%	56	4.44%	116.76%	0.68%
Más de 2.354	2,943	1.80%	49	1.66%	217.81%	3.94%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>5.60%</b>

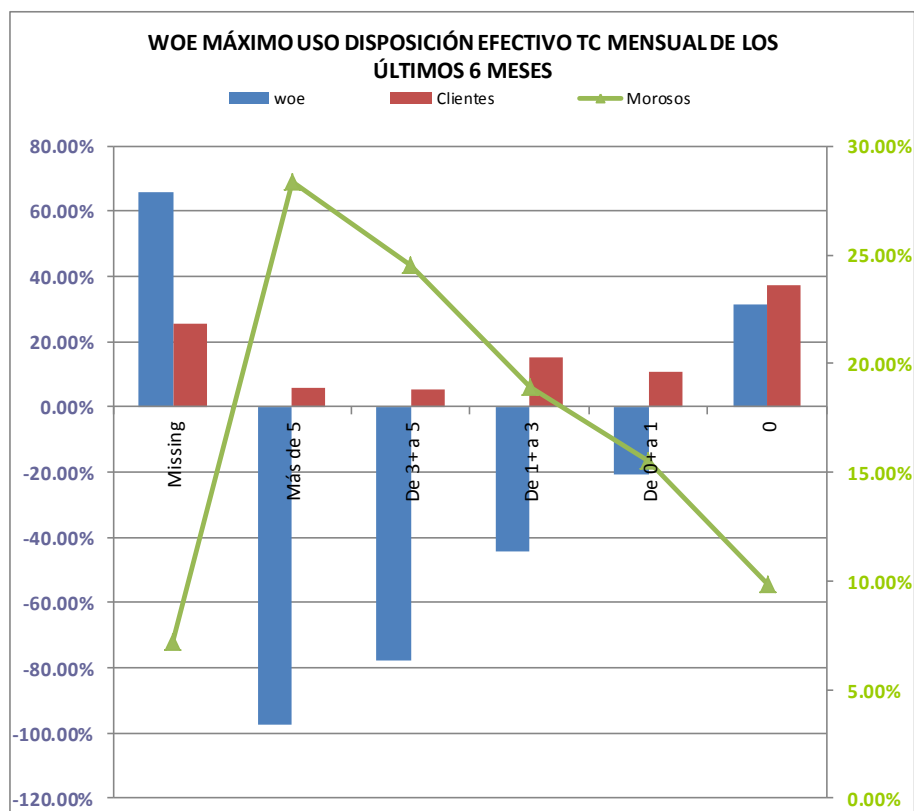
### 6.1.3.3 WOE PORCENTAJE MEDIA SALDO MEDIO AHORROS DE LOS ÚLTIMOS 6 MESES SOBRE DEUDA BBVA DEL ÚLTIMO MES DISPONIBLE V3



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coficiente	Coficiente
No posee deuda BBVA en el último mes disponible (RCC)	24,841	15.23%	1,905	7.67%	58.78%	4.22%
No posee ahorros	73,519	45.08%	12,006	16.33%	-26.66%	3.53%
Hasta 20,16486	46,598	28.57%	6,067	13.02%	-0.12%	0.00%
De 20,16486+ a 86,03352	7,768	4.76%	725	9.33%	37.32%	0.58%
De 86,03352+ a 408,7814	5,178	3.18%	340	6.57%	75.49%	1.36%
Más de 408,7814	5,178	3.18%	167	3.23%	150.09%	4.12%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	-	<b>13.82%</b>

## 6.1.4 Tarjeta de Crédito

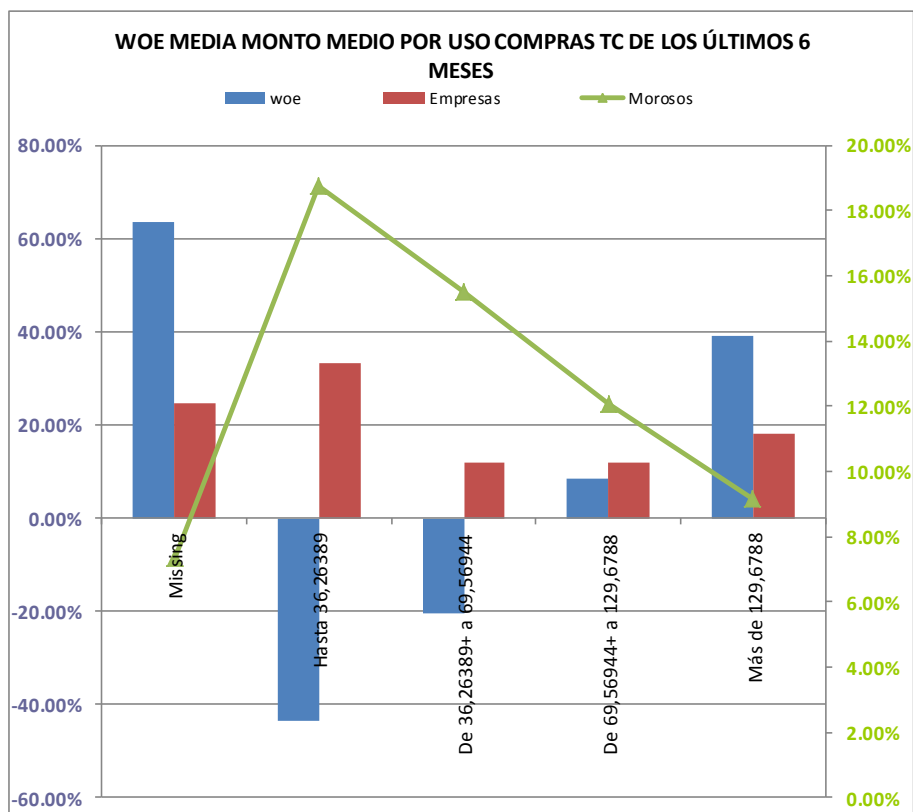
### 6.1.4.1 WOE MÁXIMO USO DISPOSICIÓN EFECTIVO TC MENSUAL DE LOS ÚLTIMOS 6 MESES



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coficiente	Coficiente
Missing	41,606	25.51%	2,989	7.18%	65.83%	8.64%
Más de 5	9,240	5.67%	2,620	28.35%	-97.35%	7.48%
De 3+ a 5	9,004	5.52%	2,210	24.54%	-77.74%	4.38%
De 1+ a 3	24,451	14.99%	4,626	18.92%	-44.52%	3.49%
De 0+ a 1	17,983	11.03%	2,794	15.54%	-20.74%	0.51%
0	60,798	37.28%	5,971	9.82%	31.68%	3.32%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	-	<b>27.83%</b>



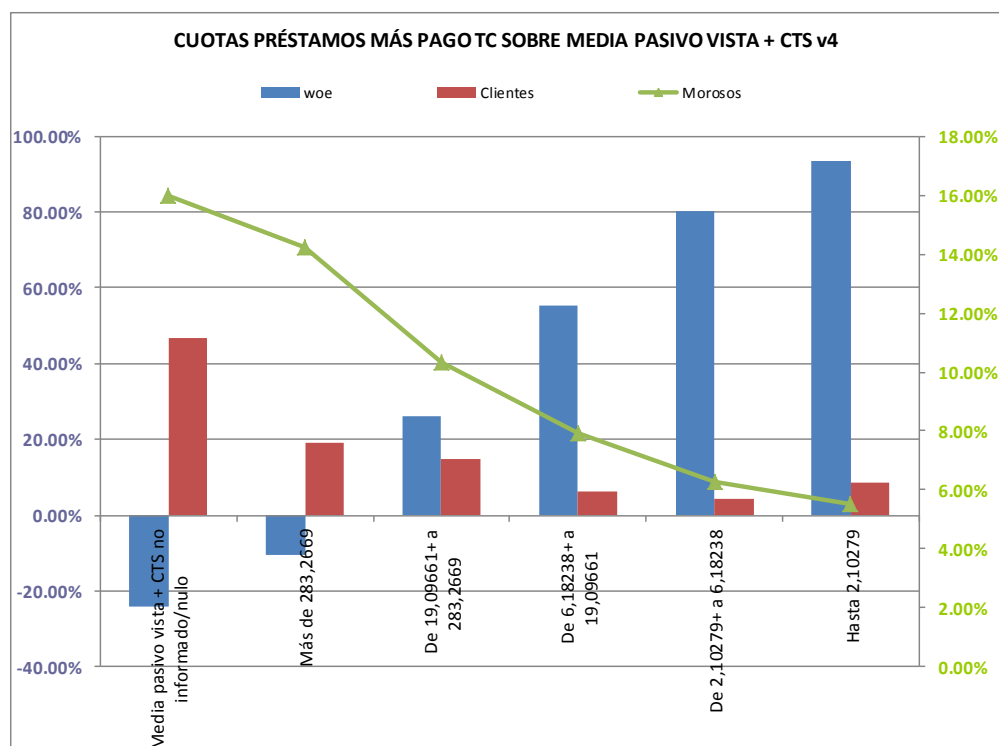
## 6.1.4.2 WOE MEDIA MONTO MEDIO POR USO COMPRAS TC MENSUAL DE LOS ÚLTIMOS 6 MESES



Valor	Empresas		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coficiente	Coficiente
Missing	40,231	24.67%	2,947	7.33%	63.73%	7.89%
Hasta 36,26389	54,055	33.15%	10,138	18.75%	-43.44%	7.32%
De 36,26389+ a 69,56944	19,654	12.05%	3,045	15.49%	-20.40%	0.54%
De 69,56944+ a 129,6788	19,658	12.05%	2,375	12.08%	8.43%	0.08%
Más de 129,6788	29,484	18.08%	2,705	9.17%	39.21%	2.40%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>18.23%</b>

## 6.1.5 Préstamos

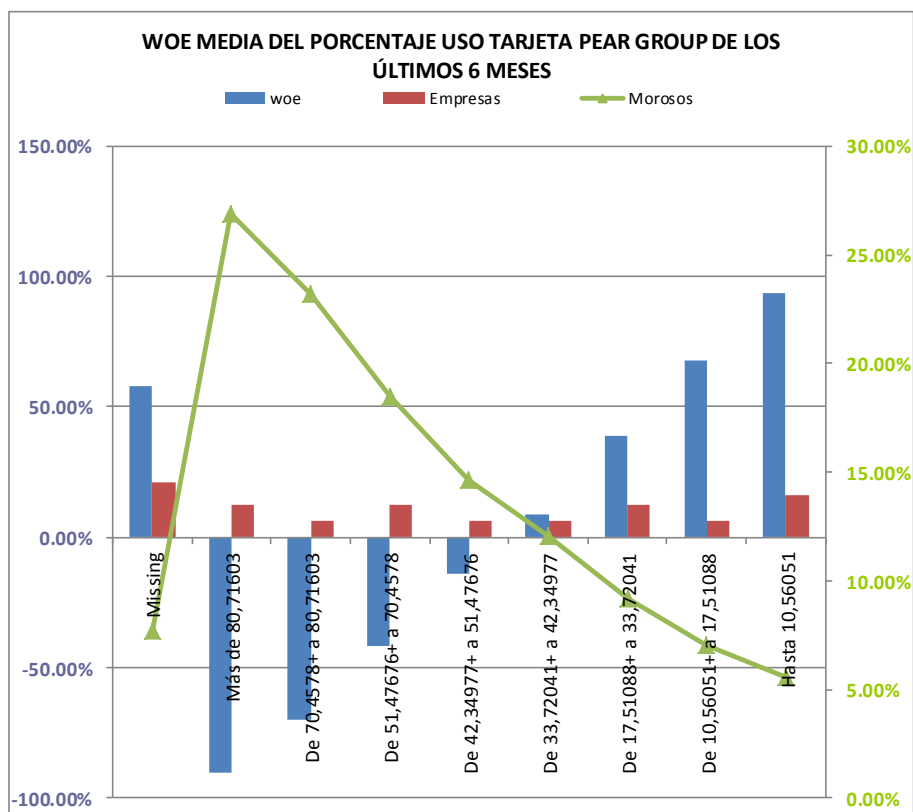
### 6.1.5.1 WOE PORCENTAJE CUOTAS PRÉSTAMOS MÁS MEDIA PAGO TC ÚLTIMOS 6 MESES SOBRE MEDIA SALDO MEDIO PASIVO VISTA + CTS V4



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coefficiente	Coefficiente
Media pasivo vista no informado/nulo	76,445	46.88%	12,234	16.00%	-24.25%	3.01%
Más de 283,2669	31,203	19.13%	4,440	14.23%	-10.41%	0.22%
De 19,09661+ a 283,2669	24,267	14.88%	2,510	10.34%	25.92%	0.91%
De 6,18238+ a 19,09661	10,402	6.38%	825	7.93%	55.13%	1.58%
De 2,10279+ a 6,18238	6,934	4.25%	435	6.27%	80.36%	2.03%
Hasta 2,10279	13,831	8.48%	766	5.54%	93.61%	5.24%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	-	<b>12.98%</b>

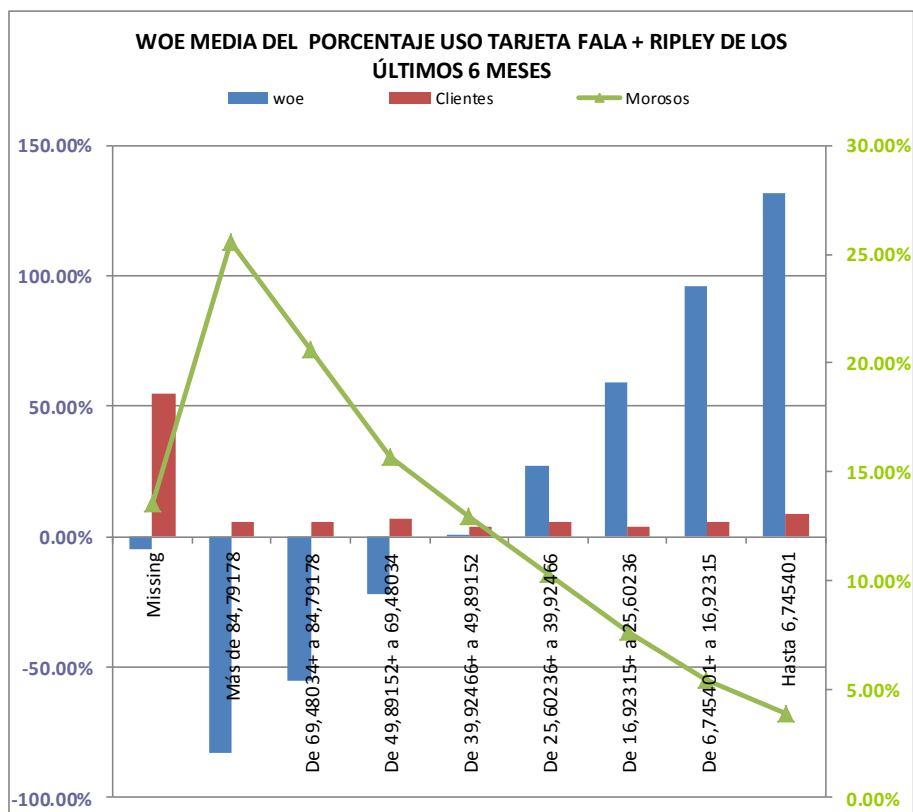
## 6.1.6 RCC

### 6.1.6.1 WOE MEDIA DEL PORCENTAJE USO TARJETA PEAR GROUP DE LOS ÚLTIMOS 6 MESES



Valor	Empresas		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coefficiente	Coefficiente
Missing	34,012	20.86%	2,628	7.73%	57.96%	5.64%
Más de 80,71603	20,652	12.66%	5,554	26.89%	-90.04%	14.00%
De 70,4578+ a 80,71603	10,324	6.33%	2,393	23.18%	-70.22%	4.00%
De 51,47676+ a 70,4578	20,652	12.66%	3,816	18.48%	-41.61%	2.55%
De 42,34977+ a 51,47676	10,326	6.33%	1,515	14.67%	-13.99%	0.13%
De 33,72041+ a 42,34977	10,326	6.33%	1,242	12.03%	8.93%	0.05%
De 17,51088+ a 33,72041	20,650	12.66%	1,902	9.21%	38.77%	1.65%
De 10,56051+ a 17,51088	10,326	6.33%	728	7.05%	67.86%	2.26%
Hasta 10,56051	25,814	15.83%	1,432	5.55%	93.43%	9.75%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>40.02%</b>

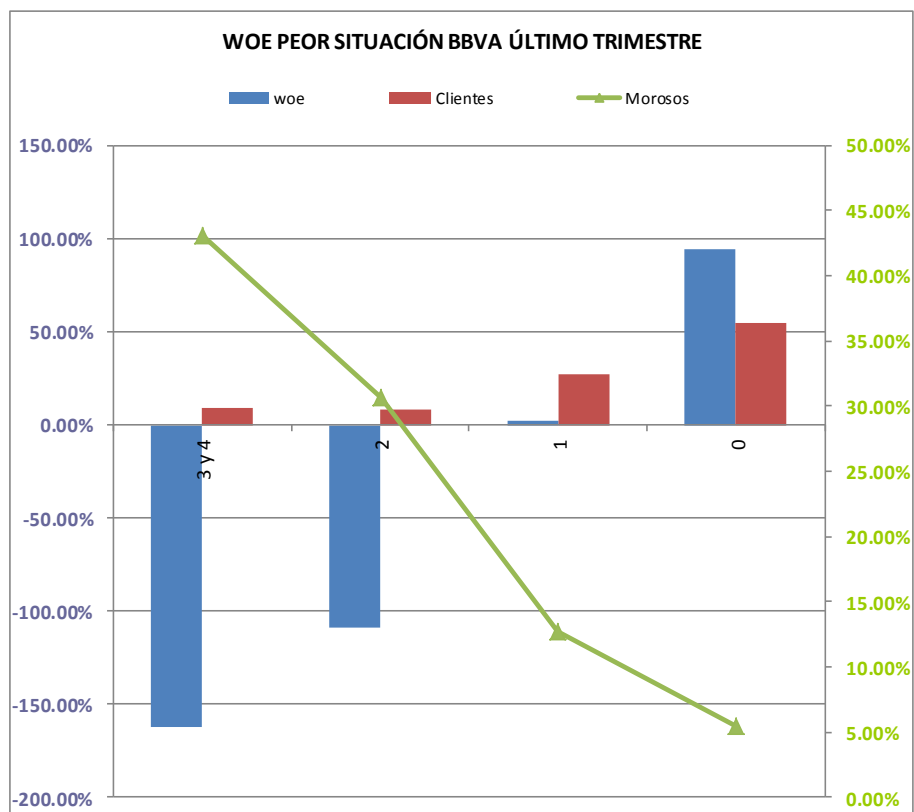
## 6.1.6.2 WOE MEDIA PORCENTAJE USO TARJETA FALABELLA + RIPLEY DE LOS ÚLTIMOS 6 MESES



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coficiente	Coficiente
Missing	89,904	55.13%	12,183	13.55%	-4.74%	0.13%
Más de 84,79178	8,781	5.38%	2,244	25.56%	-83.12%	4.96%
De 69,48034+ a 84,79178	8,781	5.38%	1,808	20.59%	-55.06%	1.99%
De 49,89152+ a 69,48034	11,709	7.18%	1,834	15.66%	-21.69%	0.37%
De 39,92466+ a 49,89152	5,854	3.59%	756	12.91%	0.81%	0.00%
De 25,60236+ a 39,92466	8,782	5.39%	900	10.25%	26.95%	0.35%
De 16,92315+ a 25,60236	5,854	3.59%	448	7.65%	59.00%	1.00%
De 6,745401+ a 16,92315	8,781	5.38%	474	5.40%	96.32%	3.49%
Hasta 6,745401	14,636	8.97%	563	3.85%	131.83%	9.58%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>21.86%</b>

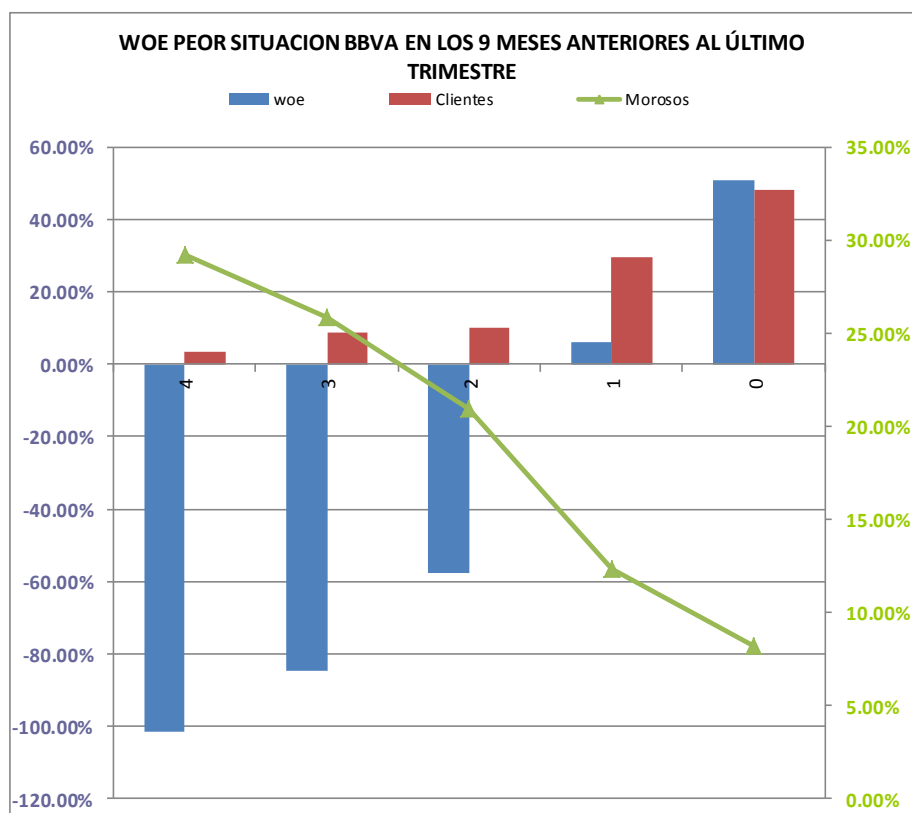
## 6.1.7 Comportamiento

### 6.1.7.1 WOE PEOR SITUACIÓN BBVA DEL ÚLTIMO TRIMESTRE



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coeficiente	Coeficiente
3 y 4	14,850	9.11%	6,395	43.06%	-162.12%	39.22%
2	13,457	8.25%	4,132	30.71%	-108.65%	14.03%
1	44,944	27.56%	5,732	12.75%	2.25%	0.01%
0	89,831	55.08%	4,951	5.51%	94.12%	34.34%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>87.60%</b>

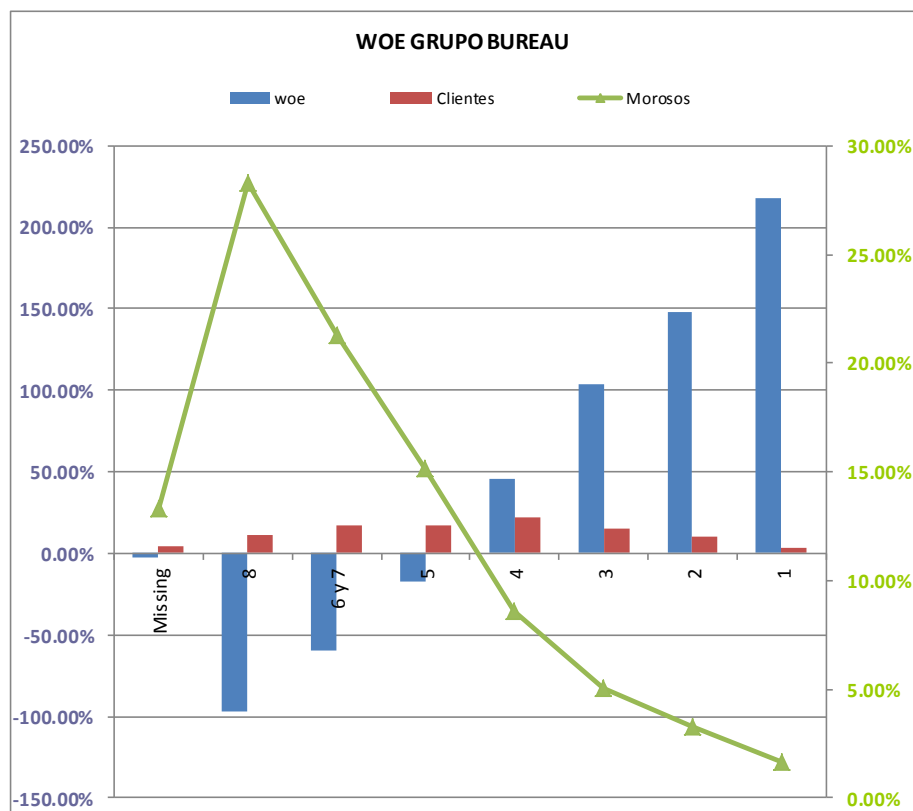
## 6.1.7.2 WOE PEOR SITUACIÓN BBVA DE LOS 9 MESES ANTERIORES AL ÚLTIMO TRIMESTRE



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coeficiente	Coeficiente
4	5,429	3.33%	1,585	29.20%	-101.45%	4.83%
3	14,608	8.96%	3,781	25.88%	-84.84%	8.65%
2	16,189	9.93%	3,398	20.99%	-57.49%	4.03%
1	48,566	29.78%	5,999	12.35%	5.90%	0.10%
0	78,290	48.01%	6,447	8.23%	51.04%	10.33%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>27.94%</b>

## 6.1.8 BURÓ

### 6.1.8.1 WOE GRUPO BURÓ



Valor	Clientes		Morosos		WOE	Inf Value
	Número	Porcentaje	Número	Porcentaje	Coeficiente	Coeficiente
Missing	6,872	4.21%	914	13.30%	-2.58%	0.00%
8	17,702	10.85%	5,004	28.27%	-96.92%	14.19%
6 y 7	28,519	17.49%	6,068	21.28%	-59.21%	7.57%
5	28,461	17.45%	4,307	15.13%	-17.62%	0.58%
4	35,533	21.79%	3,064	8.62%	46.01%	3.88%
3	24,113	14.79%	1,219	5.06%	103.24%	10.73%
2	16,582	10.17%	546	3.29%	147.95%	12.91%
1	5,300	3.25%	88	1.66%	218.09%	7.11%
<b>Total</b>	<b>163,082</b>	<b>100.00%</b>	<b>21,210</b>	<b>13.01%</b>	<b>-</b>	<b>56.97%</b>

## 6.2 Reporte técnico del modelo integración Proactivo + Grupo Riesgo Buró

### Características de la muestra de desarrollo

Periodo. desarrollo	Muestra	Fecha inicio:	201301
		Fecha fin:	201307

Tamaño. desarrollo:	Muestra	163,082.0
Tamaño muestra (Y=1):		21,210.0
Tamaño muestra (Y=0):		141,872.0

Muestra ponderada? :	entrenamiento	no
Nombre de la variable de ponderación :		binary logit
Modelo Ajustado :		0
Variable respuesta:		mora_BBVA1

Tabla 2. Likelihood & Akaike Information Criterion. Muestra Desarrollo.

Likelihood	-51806.1
AIC	103616.2

Tabla 3. Significación Global del Modelo.

Estadístico	393.4
Degreed of Freedom	8
P-Value	0.000



**Tabla 4. Contribución de las variables en la variabilidad del score. Muestra Desarrollo.**

Variable	VAR(**)	Contribucion	Deviation(*)	PercDeviation
scoreTotal	2.04	0.00	0.00	0.00
w4_edad_2	1.91	6.17	0.21	7.14
w4_antig_cliente_2	1.93	5.29	0.13	4.43
w4_med_opmon	1.92	5.72	0.20	6.58
w4_med_muso_ctc	1.84	9.43	0.20	6.88
w4_max_ue_2_tc	1.80	11.50	0.20	6.58
w4_med_pasnv	1.91	6.42	0.15	5.17
w4_ahordeu_v3	1.86	8.43	0.16	5.44
w4_med_ccte	1.97	3.45	0.11	3.69
w4_mean_tjfalrip	1.75	14.28	0.28	9.28
w4_mean_conctjpear	1.48	27.44	0.37	12.40
w4_riesgo_pasv4	1.95	4.27	0.08	2.64
w4_peor_sitbbva3m	1.44	29.23	0.42	14.05
w4_peor_sitbbva9m	1.67	17.93	0.28	9.36
w4_cod_crriesgo	1.76	13.65	0.19	6.36

(\*\*) VAR: Varianza del score sin la variable; (\*) Desviación estándar del score de la variable

**Tabla 5. Análisis de multicolinealidad de las variables del modelo. Muestra desarrollo.**

Variable	Label	VIF(**)
w4_edad_2	WOE EDAD	1.22723
w4_antig_cliente_2	WOE ANTIGÜEDAD DEL CLIENTE	1.3326
w4_med_opmon	WOE MEDIA NÚMERO DE OPERACIONES MONETARIAS DE LOS ÚLTIMOS 6 MESES	1.16076
w4_med_muso_ctc	WOE MEDIA MONTO MEDIO POR USO COMPRAS TC DE LOS ÚLTIMOS 6 MESES	1.51123
w4_max_ue_2_tc	WOE MÁXIMO USO DISPOSICIÓN EFECTIVO TC MENSUAL DE LOS ÚLTIMOS 6 MESES	1.66363
w4_med_pasnv	WOE MEDIA DEL SALDO MEDIO INVERSIÓN DE LOS ÚLTIMOS 6 MESES	1.1991
w4_ahordeu_v3	WOE MEDIA DEL SALDO MEDIO AHORROS DE LOS ÚLTIMOS 6 MESES SOBRE DEUDA BBVA ÚLTIMO MES DISPONIBLE V3	1.80453
w4_med_ccte	WOE MEDIA DEL SALDO MEDIO CUENTA CORRIENTE DE LOS ÚLTIMOS 6 MESES	1.15926
w4_mean_tjfalrip	WOE MEDIA DEL PORCENTAJE USO TARJETA FALA + RIPLEY DE LOS ÚLTIMOS 6 MESES	1.13343
w4_mean_conctjpear	WOE MEDIA DEL PORCENTAJE USO TARJETA PEAR GROUP DE LOS ÚLTIMOS 6 MESES	1.48488
W4_riesgo_pasv4	CUOTAS PRÉSTAMOS MÁS PAGO TC SOBRE MEDIA PASIVO VISTA v4	1.83053
w4_peor_sitbbva3m	WOE PEOR SITUACIÓN BBVA ÚLTIMO TRIMESTRE	1.51169
w4_peor_sitbbva9m	WOE PEOR SITUACION BBVA DE LOS 9 MESES ANTERIORES AL ÚLTIMO TRIMESTRE	1.49952
w4_cod_crriesgo	WOE GRUPO BURÓ	1.35028

(\*\*) Factor de inflación de la varianza

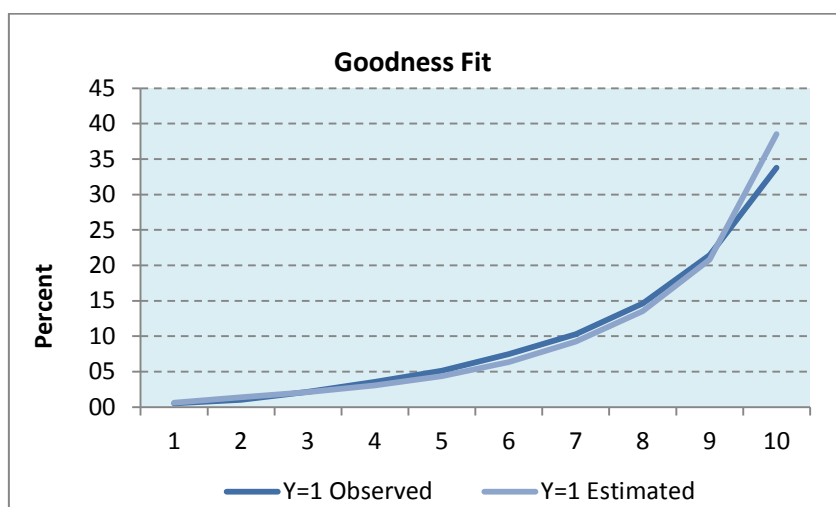
**Tabla 6. Bondad ajuste del modelo. Muestra desarrollo.**

Rango	MinProb	MaxProb	Y=0 Observed	Y=1 Observed	Y=0 estimada	Y=1 estimada
1	0.00	0.01	16,202	106	16,174	134
2	0.01	0.02	16,088	220	16,013	295
3	0.02	0.03	15,855	453	15,861	447
4	0.03	0.05	15,554	755	15,665	644
5	0.05	0.07	15,220	1,088	15,379	929
6	0.07	0.10	14,719	1,589	14,961	1,347
7	0.10	0.14	14,129	2,180	14,347	1,962
8	0.14	0.21	13,209	3,099	13,439	2,869
9	0.21	0.35	11,756	4,552	11,891	4,417
10	0.35	0.86	9,140	7,168	8,142	8,166

**Tabla 7. Contraste bondad ajuste. Muestra desarrollo.**

Hosmer- Lemeshow	Grados de libertad	P-valor
421.44	8	0

**Gráfico 1. Bondad ajuste. Muestra desarrollo.**

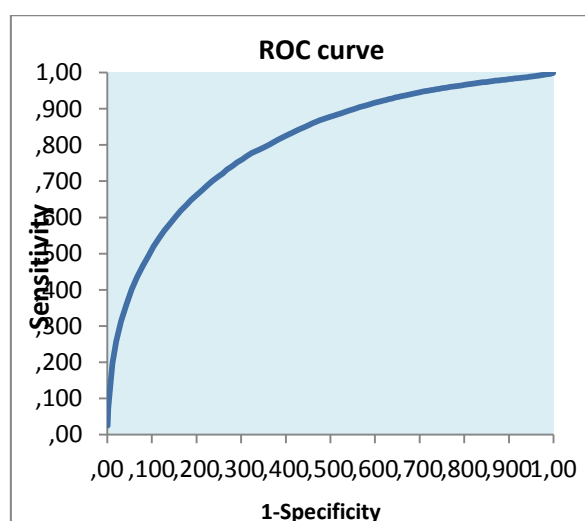


**Tabla 8. Medidas de discriminación del modelo. Muestra desarrollo.**

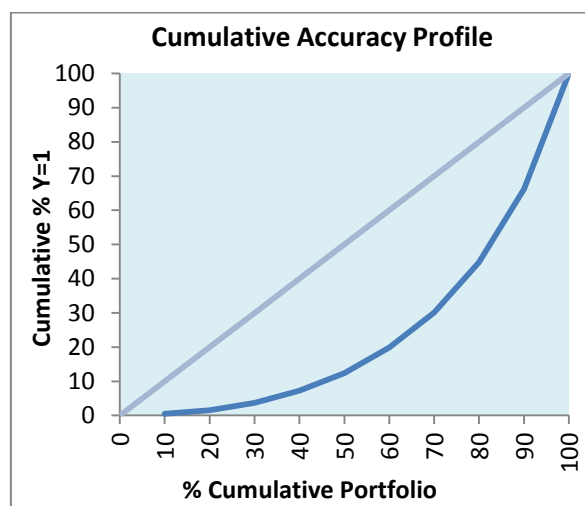
CAP (Cumulative accuracy profile) (*)	80.54
AR (Accuracy ratio) (*)	61.09
AUROC (**)	80.62
Kolmogorov-Smirnov (**)	46.50
Pietra Coefficient (**)	16.44
Somers D (**)	61.23
Bayesian error rate (**)	11.40
Kullback-Leibler Distance (**)	-4.11
CIER (**)	-24.49

(\*) calculado con 20 tramos; (\*\*) calculado con 200 tramos

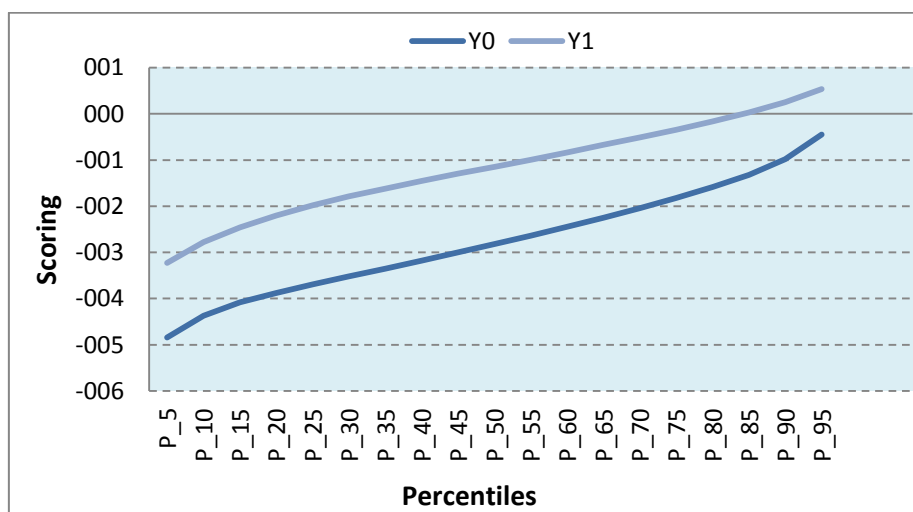
**Gráfico 2. Curva ROC. Muestra desarrollo.**



**Gráfico 3. Cumulative Accuracy profile. Muestra desarrollo.**



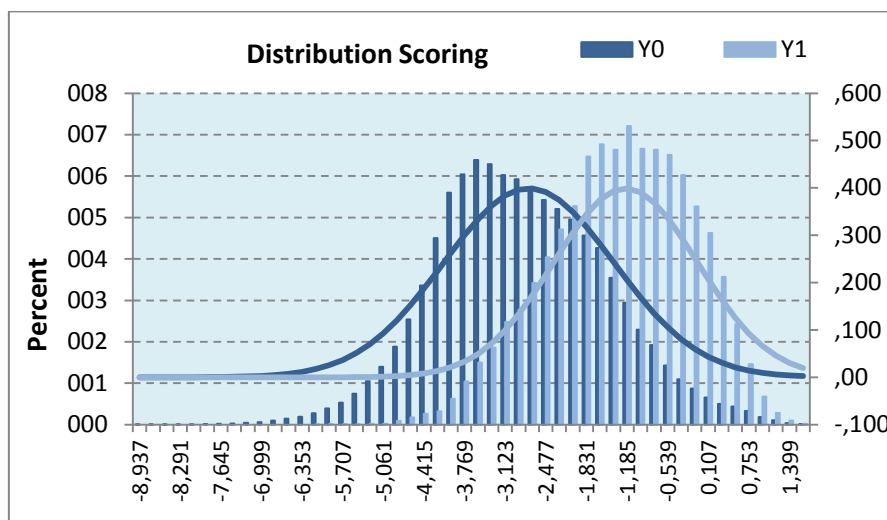
**Grafico 4. Distribución Scoring. Percentiles. Muestra desarrollo.**



**Tabla 9. Percentiles del Scoring. Muestra desarrollo.**

Percentiles	Y=0	Y=1
0%	-8.937	-6.018
5%	-4.846	-3.228
10%	-4.372	-2.781
15%	-4.083	-2.458
20%	-3.879	-2.200
25%	-3.690	-1.980
30%	-3.518	-1.784
35%	-3.350	-1.619
40%	-3.181	-1.457
45%	-3.003	-1.296
50%	-2.823	-1.147
55%	-2.639	-0.996
60%	-2.447	-0.835
65%	-2.248	-0.672
70%	-2.041	-0.509
75%	-1.823	-0.342
80%	-1.588	-0.167
85%	-1.321	0.028
90%	-0.982	0.248
95%	-0.446	0.535
100%	1.829	1.795

**Gráfico 5. Distribución Scoring. Histograma. Muestra desarrollo.**



**Tabla 10. Medidas de la distribución del Scoring. Muestra desarrollo.**

	Media	Desviación estandar	CV	Mínimo	Maximum
<b>Y=0</b>	-2.747	1.355	0.49	-8.937	1.829
<b>Y=1</b>	-1.208	1.148	0.95	-6.018	1.795

(\*) CV= coeficiente de variación

**Tabla 11. Tabla de eficiencia. Muestra desarrollo.**

Rango	Min Prob	Max Prob	Y=0	Y=1	Total	Acum Y=0	Acum Y=1	Acum Total	% Y=1 Group	% Y=1 Assum	% Y=0 Assum	% Y=1 Assum
5%	0	0.009	8128	26	8154	8128	26	8154	0.32%	0.32%	5.73%	0.12%
10%	0.009	0.014	8074	80	8154	16202	106	16308	0.98%	0.65%	11.42%	0.50%
15%	0.014	0.018	8067	88	8155	24269	194	24463	1.08%	0.79%	17.11%	0.91%
20%	0.018	0.022	8021	132	8153	32290	326	32616	1.62%	1.00%	22.76%	1.54%
25%	0.022	0.027	7953	201	8154	40243	527	40770	2.47%	1.29%	28.37%	2.48%
30%	0.027	0.033	7902	252	8154	48145	779	48924	3.09%	1.59%	33.94%	3.67%
35%	0.033	0.039	7818	337	8155	55963	1116	57079	4.13%	1.96%	39.45%	5.26%
40%	0.039	0.047	7736	418	8154	63699	1534	65233	5.13%	2.35%	44.90%	7.23%
45%	0.047	0.057	7649	505	8154	71348	2039	73387	6.19%	2.78%	50.29%	9.61%
50%	0.057	0.068	7571	583	8154	78919	2622	81541	7.15%	3.22%	55.63%	12.36%
55%	0.068	0.082	7424	730	8154	86343	3352	89695	8.95%	3.74%	60.86%	15.80%
60%	0.082	0.099	7295	859	8154	93638	4211	97849	10.53%	4.30%	66.00%	19.85%
65%	0.099	0.119	7152	1002	8154	100790	5213	106003	12.29%	4.92%	71.04%	24.58%
70%	0.119	0.144	6977	1178	8155	107767	6391	114158	14.45%	5.60%	75.96%	30.13%
75%	0.144	0.175	6666	1488	8154	114433	7879	122312	18.25%	6.44%	80.66%	37.15%
80%	0.175	0.213	6543	1611	8154	120976	9490	130466	19.76%	7.27%	85.27%	44.74%
85%	0.213	0.266	6123	2031	8154	127099	11521	138620	24.91%	8.31%	89.59%	54.32%
90%	0.266	0.347	5633	2521	8154	132732	14042	146774	30.92%	9.57%	93.56%	66.20%
95%	0.347	0.477	4782	3372	8154	137514	17414	154928	41.35%	11.24%	96.93%	82.10%
100%	0.477	0.862	4358	3796	8154	141872	21210	163082	46.55%	13.01%	100.00%	100.00%
<b>Total</b>			<b>141872</b>	<b>21210</b>	<b>163082</b>						<b>GINI</b>	<b>61.09%</b>

Tabla 12. Correlaciones.

Pearson Correlation Coefficients														
	w4_edad_2	w4_ANTIG_CLIENTE_2	W4_MED_OPMON	w4_MED_MUSOCTC	w4_MAX_UE_2_TC	w4_MED_PASNV	W4_AHORDEU_v3	w4_MED_CCTE	w4_MEAN_TJFALRIP	w4_MEAN_CONCTJPEAR	W4_riesgo_pasv4	W4_peor_sitBBVA3m	W4_peor_sitBBVA9m	w4_COD_CRRIESGO
w4_edad_2	1.000	0.383	-0.089	0.048	0.060	0.111	-0.035	0.110	0.079	0.100	-0.019	0.075	0.024	0.086
w4_ANTIG_CLIENTE_2	0.383	1.000	0.107	0.115	0.138	0.249	0.101	0.200	0.137	0.110	0.187	0.033	-0.055	0.117
W4_MED_OPMON	-0.089	0.107	1.000	0.032	-0.002	0.179	0.208	0.209	0.067	0.009	0.282	0.079	0.071	0.093
w4_MED_MUSOCTC	0.048	0.115	0.032	1.000	0.552	0.030	-0.059	0.060	0.075	0.209	0.007	0.120	0.118	0.238
w4_MAX_UE_2_TC	0.060	0.138	-0.002	0.552	1.000	0.052	0.087	0.019	0.147	0.388	0.140	0.145	0.101	0.265
w4_MED_PASNV	0.111	0.249	0.179	0.030	0.052	1.000	0.247	0.250	0.114	0.121	0.273	0.082	0.064	0.124
W4_AHORDEU_v3	-0.035	0.101	0.208	-0.059	0.087	0.247	1.000	0.060	0.126	0.253	0.625	0.230	0.196	0.135
w4_MED_CCTE	0.110	0.200	0.209	0.060	0.019	0.250	0.060	1.000	0.079	0.061	0.203	0.051	0.026	0.118
w4_MEAN_TJFALRIP	0.079	0.137	0.067	0.075	0.147	0.114	0.126	0.079	1.000	0.305	0.120	0.138	0.112	0.207
w4_MEAN_CONCTJPEAR	0.100	0.110	0.009	0.209	0.388	0.121	0.253	0.061	0.305	1.000	0.185	0.263	0.242	0.394
W4_riesgo_pasv4	-0.019	0.187	0.282	0.007	0.140	0.273	0.625	0.203	0.120	0.185	1.000	0.173	0.148	0.150
W4_peor_sitBBVA3m	0.075	0.033	0.079	0.120	0.145	0.082	0.230	0.051	0.138	0.263	0.173	1.000	0.540	0.318
W4_peor_sitBBVA9m	0.024	-0.055	0.071	0.118	0.101	0.064	0.196	0.026	0.112	0.242	0.148	0.540	1.000	0.311
w4_COD_CRRIESGO	0.086	0.117	0.093	0.238	0.265	0.124	0.135	0.118	0.207	0.394	0.150	0.318	0.311	1.000